Sepp Hochreiter



Von Europas KI-Pionier

WAS KANN KÜNSTLICHE INTELLIGENZ?

Wird sie Wohlstand für alle schaffen, Krebs heilen und das Klimaproblem lösen?

ecowing

Sepp Hochreiter

WAS KANN KÜNSTLICHE INTELLIGENZ?

Sepp Hochreiter

WAS KANN KÜNSTLICHE INTELLIGENZ?

Wird sie Wohlstand für alle schaffen, Krebs heilen und das Klimaproblem lösen?

ecowing



Sämtliche Angaben in diesem Werk erfolgen trotz sorgfältiger Bearbeitung ohne Gewähr.
Eine Haftung der Autoren beziehungsweise Herausgeber und des Verlages ist ausgeschlossen.

1. Auflage
© 2025 ecoWing Verlag bei Benevento Publishing Salzburg – Wien, eine Marke der
Red Bull Media House GmbH, Wals bei Salzburg

Alle Rechte vorbehalten, insbesondere das des öffentlichen Vortrags, der Übertragung durch Rundfunk und Fernsehen sowie der Übersetzung, auch einzelner Teile. Kein Teil des Werkes darf in irgendeiner Form (durch Fotografie, Mikrofilm oder andere Verfahren) ohne schriftliche Genehmigung des Verlages reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Medieninhaber, Verleger und Herausgeber:
Red Bull Media House GmbH
Oberst-Lepperdinger-Straße 11–15; 5071 Wals bei Salzburg, Österreich
info@at.redbullmediahouse.com

Layout und Satz: MEDIA DESIGN: RIZNER.AT
Gesetzt aus der Palatino, Resolve

Umschlagmotiv: © Nolan / stock.adobe.com, generiert mit KI
Umschlaggestaltung: b3K design, Andrea Schneider, diceindustries
Autorenillustration: Claudia Meitert / carolineseidler.com

Printed by PNB Print in Lettland
ISBN: 978-3-7110-5382-4
e-ISBN: 978-3-7110-0365-2

Inhaltsverzeichnis

Einleitung

1 Lernende Netze

2 KI mit Gedächtnis

3 Mensch versus Maschine

4 Was kommt nach ChatGPT?

5 Die Industrialisierung der KI

6 Die Welt simulieren

7 Digitale Zwillinge

8 Komplexität beherrschen

9 Chance für Europa

10 KI kontrollieren

11 Warum wir KI brauchen

GLOSSAR

Quellenhinweise

Einleitung

Als ich 1991 meine Diplomarbeit über »dynamische neuronale Netze« schrieb, skizzierte ich darin Methoden, die erst viele Jahrzehnte später ihre Anwendungen finden sollten: Computer, die Musik komponieren, künstliche Agenten, die selbstständig lernen, und Systeme, die komplexe Prozesse steuern.

Künstliche Intelligenz erlebt heute einen Boom, den ich selbst nicht für möglich gehalten hätte. Zugleich polarisiert keine Technologie so stark wie die KI. Während die einen grenzenlose Möglichkeiten sehen, warnen die anderen vor unkalkulierbaren Risiken. KI soll gewaltige Produktivitätsfortschritte bringen, das Klima retten und Krebs heilen. Aber was kann die Technologie wirklich? Stehen wir tatsächlich vor einer neuen industriellen Revolution – oder erleben wir nur einen Hype, der bald wieder verpuffen wird?

Dieses Buch handelt von meinem wechselvollen Weg in der KI-Forschung, es geht um Erfolge und Niederlagen, um Durchbrüche und Durststrecken. Vor allem aber geht es um die Möglichkeiten und Grenzen der KI.

Ein großer Teil der öffentlichen Diskussion wird durch überhöhte Versprechungen angetrieben. Es gibt Stimmen – etwa von Sam Altman, dem CEO von OpenAI –, die behaupten, wir stünden kurz davor, eine AGI (Artificial General Intelligence) zu erreichen. Das halte ich für übertrieben. Kein KI-System hat bisher die Breite menschlicher Fähigkeiten erreicht. Kein KI-System fährt autonom ein Auto, macht die Wäsche oder löst unsere Alltagsprobleme. Die Realität ist, dass KI sehr gut in spezifischen Bereichen arbeitet, wie etwa in der Verarbeitung und Reproduktion von Wissen. Aber von einer »allgemeinen Intelligenz« oder gar einer

»Superintelligenz« sind wir weit entfernt. Ganz abgesehen davon, dass die Definition dieser Begriffe unklar ist.

Warum also solche Behauptungen? Möglicherweise spielen hier wirtschaftliche Interessen eine Rolle, etwa die Steigerung des Marktwerts von Unternehmen oder das Einwerben von Forschungsgeldern. Gleichzeitig gibt es Stimmen von Kritikern, die auf die offensichtlichen Schwächen aktueller KI-Modelle hinweisen, wie ihre Unfähigkeit, logisch zu folgern, sich auf neue Situationen einzustellen oder einen Kontext konsistent zu verarbeiten.

Die Entwicklung der KI steht derzeit an einem Wendepunkt. Seit der Veröffentlichung von ChatGPT haben vor allem amerikanische Technologiekonzerne Milliarden von Dollar in immer größere und leistungsfähigere Textmodelle investiert. Das von US-Präsident Donald Trump angekündigte Infrastrukturprojekt »Stargate« soll in den nächsten Jahren weitere 500 Milliarden Dollar in zusätzliche Rechenkapazitäten pumpen. Doch die Zweifel wachsen, ob dies der richtige Weg ist.

Da die großen Internetfirmen derzeit immer größere KI-Modelle vorstellen, könnte man meinen, dass Intelligenz aus großer Rechenleistung und großen Datenmengen entsteht. Die Realität ist jedoch anders. Die Schwächen aktueller großer KI-Systeme zeigen sich besonders in logischen Aufgaben. Ein oft zitiertes Beispiel ist das Problem mit Pfirsichen: Ein KI-Modell berechnet korrekt, wie viele Pfirsiche übrigbleiben, wenn man welche hinzufügt oder entfernt. Doch wenn zusätzliche, irrelevante Informationen eingebaut werden – etwa das Gewicht oder die Farbe der Pfirsiche –, scheitert das System. Menschen erkennen sofort, dass diese Details für die Aufgabe keine Rolle spielen. Ein KI-System hingegen verliert den Fokus und liefert falsche Ergebnisse.

Dieses Beispiel zeigt, dass KI zwar beeindruckend erscheinen kann, aber immer noch grundlegende Verständnisprobleme hat. So imitieren große Textmodelle Intelligenz, indem sie Texte reproduzieren und kombinieren, die sie während des Trainings gelernt haben. Sie »verstehen« jedoch nicht im menschlichen Sinne. Letztlich sind sie nichts anderes als Datenbanken,

die riesige Textmengen wie Papageien auswendig gelernt haben. Das gilt auch für die Modelle des chinesischen Start-ups DeepSeek, die zuletzt viel Staub aufgewirbelt haben.

So beeindruckend die Textgenerierung auch ist: Was mich viel mehr interessiert, ist eine KI, die unsere Welt versteht und modellieren kann. Eine KI, die physikalische und biologische Prozesse nachbilden kann, um neue Medikamente zu finden, den Klimawandel zu bekämpfen oder komplexe industrielle Systeme zu optimieren.

Auf dem Weg zur Maschinenintelligenz fehlt meiner Überzeugung nach etwas Grundlegendes – eine entscheidende neue Idee. Genau darum geht es mir in diesem Buch. Wir müssen besser verstehen, wie wir Menschen aus wenigen Beispielen lernen. Wie man die physikalische Welt versteht. Und wie man Probleme löst, die man nie zuvor gesehen hat.

Wir stehen an der Schwelle zur dritten Phase der Künstlichen Intelligenz: der Industrialisierung. Nach der Grundlagenforschung und der Skalierung ist es nun an der Zeit, KI-Methoden in die wirtschaftliche Anwendung zu überführen. Gerade für Europa bietet sich jetzt eine einmalige Gelegenheit – wenn es gelingt, die Zusammenarbeit zwischen Forschungseinrichtungen und Unternehmen zu intensivieren und Innovationen gezielt in die Praxis zu bringen. Das nötige Wissen und das industrielle Know-how haben wir.

Das wahre Potenzial von KI liegt darin, komplexe Probleme zu lösen, die unsere menschlichen Fähigkeiten übersteigen. Wir brauchen sie, um den Klimawandel zu bekämpfen, um Krebs zu heilen, um das Energie- oder Verkehrsproblem zu lösen. Künstliche Intelligenz darf kein Selbstzweck sein. Am Ende muss sie uns Menschen dienen.

1

Lernende Netze

Als 13-Jähriger stieß ich in einer Regionalzeitung auf einen Artikel über Schach. Das Spiel zog mich sofort in seinen Bann. Ich wollte alles über Schach wissen, besiegte irgendwann meinen Onkel und trat einem Schachklub bei. Schach war für mich dabei mehr als nur ein Spiel. Es schärfte nicht nur mein logisches und strategisches Denken. Durch das Schachspielen lernte ich auch, mich über lange Zeiträume hinweg auf ein Problem zu konzentrieren – eine Fähigkeit, die mir bei meiner wissenschaftlichen Karriere später entscheidend half.

Ich wuchs auf einem Bauernhof in Bayern auf. Meine Kindheit war geprägt vom Landleben, von Fußball und Skifahren – und eigentlich war mein Weg vorgezeichnet: Als Erstgeborener sollte ich den elterlichen Hof übernehmen. Der Besuch eines Gymnasiums kam nicht in Frage. Doch ich war neugierig, ich wollte lernen. Im Schachklub traf ich auf Gymnasiasten, die mir Bücher über Physik und Mathematik empfahlen. Im Selbststudium eignete ich mir das Mathematikwissen der ersten Universitätssemester an. Ohne Abitur konnte ich zwar nicht Mathematik studieren. Mein Schulabschluss ermöglichte mir aber zumindest eine Studienberechtigung für ein Informatikstudium an der FH München, das dortige Vordiplom wiederum den Zugang zu einem Informatikstudium an der TU München. Diese Mischung aus Mathematik und Technik wurde später zu meinem größten Vorteil.

Als ich Mitte der 1980er-Jahre mein Studium begann, war die Informatik noch ein junges Fach. Von künstlicher Intelligenz war noch nicht viel zu sehen, zumindest nicht in der modernen Form, wie wir sie heute kennen. Es gab erste Ansätze, neuronale Netze zu verstehen und anzuwenden, aber es

fühlte sich oft an wie Science-Fiction. Denn diese neue Disziplin versprach etwas, das über die Mathematik hinausging: Maschinen zu entwickeln, die menschliche Probleme lösen können. Die Idee, dass ein Computer nicht nur rechnen, sondern auch lernen und »denken« könnte, übte eine unwiderstehliche Faszination auf mich aus. Schon als Student spürte ich, dass hier etwas grundlegend Neues begann.

Eine neue Art zu denken

In der KI-Forschung gab es damals zwei Welten: die klassische symbolische KI – und das völlig neue, unerschlossene Feld der subsymbolischen KI, der neuronalen Netze also. Symbolische KI arbeitet mit Methoden, die auf klaren Regeln, Formeln und Logik basieren. Subsymbolische KI, wie typischerweise neuronale Netze, lernt hingegen aus Daten und generalisiert auf neue Fälle.

In Europa dominierte in den 80er-Jahren die symbolische KI. Expertensysteme, die mit symbolischem Wissen gefüttert wurden, lieferten damals durchaus beeindruckende Ergebnisse, insbesondere in spezialisierten Bereichen wie der Medizin oder der Ingenieurwissenschaft. Trotz ihrer Erfolge stieß die symbolische KI allerdings an Grenzen. Sie war stark abhängig von vorab definierten Regeln und konnte mit unstrukturierten oder unvollständigen Informationen kaum umgehen. Genau da setzte die subsymbolische KI an, die auf lernenden Systemen wie neuronalen Netzen basiert.

Die heutige KI-Forschung bewegt sich zunehmend in Richtung subsymbolischer Ansätze, da diese in vielen praktischen Anwendungen überlegen sind. Dennoch bleibt die symbolische KI relevant, insbesondere in Bereichen, die hohe Sicherheit und Genauigkeit erfordern.

Expertensysteme, Logiksysteme und andere Ansätze der symbolischen KI fand ich zwar interessant, doch mir fehlte darin die Faszination des Neuen und Unbekannten. Also stürzte ich mich auf das neue, kaum

erforschte Gebiet der neuronalen Netze, das sich von der Funktionsweise des menschlichen Gehirns inspirieren ließ. Es war eine völlig neue Herangehensweise, die nicht auf starren Regeln beruhte. Was mich besonders interessierte, war die Fähigkeit neuronaler Netze, aus Beispielen zu lernen. Ein Netz wird mit Daten gefüttert – mit Eingaben und den dazugehörigen Ausgaben – und entwickelt daraus ein Modell, das diese Beziehung beschreibt. Das eröffnete ungeahnte Möglichkeiten, Probleme zu lösen.

Neuronale Netze waren zu dieser Zeit allerdings noch ein Nischenthema. Sie galten als unzuverlässig und schwer verständlich, und es gab kaum praktische Anwendungen. Ich spürte jedoch, dass die neuronalen Netze nicht nur sehr interessant waren, sondern zudem ein großes Potenzial hatten, auch wenn sie damals noch ganz am Anfang standen.

Heute sind neuronale Netze das Herzstück moderner KI-Systeme und haben sich von einem spekulativen Forschungsgebiet zu einer Schlüsseltechnologie entwickelt. Ihre Vielseitigkeit macht sie in zahlreichen Bereichen unverzichtbar, von der Bild-, Text- und Spracherkennung bis hin zu Anwendungen in der Informationstechnologie, der Medizin und der Finanzwelt. Doch was genau macht neuronale Netze so einzigartig?

Ein wesentlicher Vorteil liegt in ihrer Fähigkeit zur Generalisierung. Im Gegensatz zu klassischen regelbasierten Systemen lernen neuronale Netze Muster und Strukturen direkt aus Daten. Diese Herangehensweise ermöglicht es ihnen, auch in komplexen, unvorhersehbaren Situationen zu bestehen, für die keine Regeln bekannt sind. Beispielsweise können sie aus Millionen von Bildern lernen, Hunde von Katzen zu unterscheiden, ohne dass ihnen explizit Regeln vorgegeben werden.

Eine Sackgasse?

Als ich begann, mich intensiv mit KI zu beschäftigen, war vieles an den neuronalen Netzen noch spekulativ. Es gab keine Gewissheit, dass sie

wirklich funktionieren würden, und viele Wissenschaftler hielten sie für eine Sackgasse. Ich erinnere mich an Diskussionen, in denen KI-Experten argumentierten, dass neuronale Netze keine Zukunft hätten. Die großen Erfolge, die wir heute mit Deep Learning feiern, waren damals kaum vorstellbar. Es war aber gerade diese Unsicherheit, die mich reizte. Etwas zu erforschen, das viele Möglichkeiten bietet, noch ungelöste Probleme hat, neue Ideen zulässt und noch nicht ausgeforscht ist.

Ich kam über Jürgen Schmidhuber, der an der TU München lehrte, mit neuronalen Netzen in Berührung. Wie gesagt, gehörte ich damals zu einer Minderheit in der KI-Forschung. Unsere Community war überschaubar; wir kannten uns fast alle persönlich. Auf den internationalen Konferenzen waren wir vielleicht 200 bis 300 Teilnehmer, während die symbolische KI die großen Bühnen bespielte. Was uns von den etablierten KI-Forschern unterschied, war nicht nur unser Forschungsthema, sondern auch unser Ansatz. Wir experimentierten, probierten neue Ideen aus. Wir lernten zu akzeptieren, dass vieles nicht funktionierte. Diese Offenheit machte die Arbeit aufregend, aber sie isolierte uns auch von der etablierten KI-Welt.

Zu neuronalen Netzen gab es damals auch wenig Literatur, kaum etablierte Methoden – wir mussten alles selbst herausfinden. Ich erinnere mich, wie ich als Student mit simplen Programmen experimentierte, die uns an der Universität zur Verfügung gestellt wurden. Wir hatten keine großen Modelle mit Milliarden von Parametern, keine Cloud-Infrastruktur, die uns unendliche Rechenleistung bot. Alles war handgemacht, von den Algorithmen bis zu den Experimenten.

Die Konferenz NeurIPS (Neural Information Processing Systems) war das Zentrum dieser Entwicklung, und obwohl sie damals kaum Teilnehmer aus Europa hatte, wurde dort die Zukunft der KI geschrieben. Heute kommen über 15.000 Menschen aus aller Welt zusammen, um die neuesten Entwicklungen zu präsentieren. Konferenzen wie NeurIPS, ICML (International Conference on Machine Learning) oder ICLR (International Conference on Learning Representations) sind der Puls der KI-Forschung – und sie zeigen, wie sehr das Feld gewachsen ist.

Die KI hat sich seither von einer Nischendisziplin zu einem globalen Phänomen entwickelt, das fast jede Branche verändert. Doch die Wurzeln liegen in den bescheidenen Anfängen, in den Tagen, als wir erste neuronale Netze bauten. Rückblickend war ich mittendrin in einer der spannendsten Phasen der KI-Geschichte: in der Zeit, als neuronale Netze zunächst ignoriert, dann belächelt und schließlich gefeiert wurden.

Der Aufbau neuronaler Netze

Neuronale Netze basieren auf einer Analogie zum menschlichen Gehirn. Dieses besteht aus Milliarden von Neuronen, die durch synaptische Verbindungen miteinander kommunizieren. Jedes Neuron kann Signale senden oder empfangen und dadurch andere Neuronen aktivieren oder hemmen. Künstliche neuronale Netze versuchen, diese Struktur nachzuahmen. Sie bestehen aus einer Vielzahl einfacher Verarbeitungseinheiten – Neuronen –, die miteinander vernetzt sind. Diese Netzwerke können durch Anpassung der synaptischen Verbindungen lernen, komplexe Aufgaben zu bewältigen.

Das einfachste neuronale Netz ist das sogenannte »Perzeptron«, das 1943 von Warren McCulloch und Walter Pitts konzipiert wurde. Es besteht aus nur einem einzigen Neuron, das externe Signale verarbeitet, ähnlich wie ein Neuron in der Retina unseres Auges. Doch unser Gehirn ist weit komplexer: Es verfügt über Neuronen, die in Netzwerken organisiert sind und miteinander Signale austauschen.

Ein neuronales Netz besteht aus Eingabeneuronen, die Informationen von außen aufnehmen, und Ausgabeneuronen, die Ergebnisse liefern. Dazwischen liegen oft mehrere Schichten, die die Verarbeitung der Daten ermöglichen. Der Unterschied zwischen verschiedenen Netzen liegt in der Stärke der Verbindungen zwischen den Neuronen. Einige Verbindungen können erregend sein, andere hemmend. Diese Unterschiede bestimmen, wie das Netz funktioniert.

Das Lernen in einem neuronalen Netz erfolgt durch Anpassung dieser Verbindungen. Stellen Sie sich vor, Sie möchten ein Netz so trainieren, dass es die mathematische Funktion $f(x)=x^2$ repräsentiert. Das heißt, wenn das Netz x an den Eingabeneuronen erhält, dann soll es x^2 an die Ausgabeneuronen liefern. Beim Training geben Sie dem Netz eine Reihe von Eingabewerten – zum Beispiel 1, 2, 3 – und die dazugehörigen Ausgaben – also 1, 4, 9. Anfangs sind die Verbindungen im Netz zufällig gesetzt, sodass das Netz falsche Ergebnisse liefert. Etwa: Es gibt für die Eingabe 2 die Ausgabe 3 statt 4. Das Netz passt daraufhin die Verbindungen an, um die Abweichung zu korrigieren. Dieser Prozess wiederholt sich, bis das Netz die korrekten Ausgaben für alle Trainingsdaten liefert.

Hat ein neuronales Netz fertig gelernt, werden die Daten, anhand derer es gelernt hat, weggenommen. Das neuronale Netz hat die wesentlichen Informationen aus den Daten in seinen Verbindungen gespeichert, so wie wir Menschen aus einem Buch lernen und dann das Buch nicht mehr verwenden, da wir die wesentlichen Dinge nun wissen.

Muster erkennen

Das Lernen neuronaler Netze unterscheidet sich grundlegend von herkömmlicher Informatik. In klassischen Computerprogrammen wird eine Lösung deduktiv entwickelt: Ein Mensch schreibt ein Programm, das auf Basis der von diesem Menschen erdachten Regeln ein Problem löst. Neuronale Netze hingegen lernen Muster und Regeln anhand von Daten basierend auf einem Lernalgorithmus. In beiden Fällen werden dann die Regeln (erdachte oder gelernte) auf neue, unbekannte Daten angewendet.

Ein Beispiel: Sie trainieren ein Netz mit Bildern von Hunden und Katzen. Jedes Bild ist mit einem Label versehen – »Hund« oder »Katze«. Das Netz passt seine Verbindungen so an, dass es anhand der Pixelmuster in den Bildern lernt, Hunde von Katzen zu unterscheiden. Nach dem Training soll das Netz auch Bilder korrekt klassifizieren, die es nie zuvor gesehen

hat. Das ist das Prinzip der Generalisierung, das neuronale Netze so mächtig macht und das mit erdachten Regeln nicht erreicht wurde.

Doch dieser Ansatz hat seine Tücken. Ein Netz könnte statt der zugrunde liegenden Regeln nur die Trainingsdaten auswendig lernen – ein Phänomen, das als »Overfitting« bezeichnet wird. Ein überangepasstes Netz liefert zwar perfekte Ergebnisse für die Trainingsdaten, versagt jedoch bei neuen Eingaben. Um das zu vermeiden, setzt man spezielle Techniken ein, etwa die Nutzung von Regularisierungsmethoden, Validierungsdatensätzen oder die Begrenzung der Netzkomplexität. Die Techniken helfen dabei, dass ein KI-Modell auch mit neuen, unbekannten Daten umgehen kann und nicht nur die Trainingsdaten auswendig lernt.

Ein weiteres Problem ist die fehlende Garantierbarkeit. Anders als symbolische KI, die auf Logik und klaren Regeln basiert, bietet die subsymbolische KI keine vollständige Kontrolle. Das Netz könnte für bestimmte ungesehene Eingaben unerwartet falsche Ergebnisse liefern, und es gibt keine Möglichkeit, dies im Voraus sicher auszuschließen. Ein ungelöstes Problem ist bis heute die Black-Box-Natur der neuronalen Netze. Anders als bei symbolischer KI ist es schwierig zu erklären, warum ein Netz eine bestimmte Entscheidung trifft. Dieses Problem der Interpretierbarkeit ist gerade bei sicherheitskritischen Anwendungen sehr relevant. Denn wie kann man sicherstellen, dass ein KI-System die richtigen Gründe für eine Entscheidung hat?

Die Vielseitigkeit neuronaler Netze zeigt sich bei vielen modernen Anwendungen. Sie erkennen Gesichter auf Fotos, übersetzen Texte in Echtzeit oder analysieren komplexe Datenmuster in Finanzmärkten. Diese Erfolge basieren auf der Fähigkeit neuronaler Netze, generalisieren zu können. Doch die Generalisierung ist keine Selbstverständlichkeit. Ein Netz könnte sich an Mustern orientieren, die zwar mit der Datenerzeugung korreliert sind, aber nicht mit der Aufgabenstellung. Das Netz sollte allgemeine Muster der Aufgabenstellung in den Daten erkennen, die auch bei neuen Daten vorhanden sind.

Bleiben wir hierzu bei unserem Katzen-/Hunde-Beispiel: Wenn ein Netz Bilder von Katzen und Hunden unterscheiden soll, könnte es lernen, dass ein grüner Hintergrund häufiger bei Hundebildern vorkommt, anstatt die tatsächlichen Merkmale von Katzen und Hunden zu analysieren. Hunde werden öfter im Freien fotografiert als Katzen. Solche Fälle erfordern ein sorgfältiges Design der Lernmethoden und Tests nach dem Lernen, um sicherzustellen, dass das Netz Merkmale lernt, die generalisieren, die also bei neuen Bildern auch noch vorhanden sind.

Aus Fehlern lernen

Neuronale Netze haben in den letzten Jahrzehnten eine beeindruckende Entwicklung durchlaufen. Mit dem Aufkommen des Deep Learning wurden sie leistungsfähiger, vielseitiger und in vielen Bereichen unverzichtbar. Doch auf dem Weg dorthin waren einige technische und mathematische Probleme zu lösen. Einer der Schlüsselaspekte dabei ist, wie die sogenannten »Gewichte« und die Neuronen in den verschiedenen Schichten in einem neuronalen Netz zusammenarbeiten, um die gewünschten Ergebnisse zu liefern.

Heutige neuronale Netze arbeiten mit Millionen oder sogar Milliarden von Parametern, den genannten »Gewichten«, die für das Lernen der Netze entscheidend sind. Gewichte repräsentieren die Stärke der Verbindung zwischen Neuronen. Jede Verbindung trägt dazu bei, ein Signal von einem Neuron zum nächsten weiterzugeben. Die Gewichte funktionieren in gewisser Weise ähnlich wie Drehregler, die zum Beispiel die Lichtstärke regulieren. Ein höheres Gewicht bedeutet, dass das Signal verstärkt wird, während ein niedrigeres Gewicht das Signal schwächt.

Das Lernen der neuronalen Netze ist ein iterativer, also ein sich wiederholender Prozess, der sich auf die Anpassung der Gewichte konzentriert. Zu Beginn des Trainings sind diese Gewichte zufällig gewählt, und das Netz liefert zumeist fehlerhafte Ausgaben. Während des

Trainingsprozesses passt ein Lernalgorithmus die Gewichte an, um die Fehler in den Vorhersagen zu minimieren, damit das Netz die gewünschte Aufgabe besser erfüllt.

Eine zentrale Lernmethode ist der Backpropagation-Algorithmus. Dieses Verfahren – schon 1970 von dem Finnen Seppo Linnainmaa beschrieben, dann in den 1980er-Jahren von Forschern wie David Rumelhart und Geoffrey Hinton popularisiert – ermöglicht es, neuronale Netze effizient zu trainieren. Der Backpropagation-Algorithmus analysiert dabei, wie stark sich die Ausgabe des Netzes vom gewünschten Ergebnis unterscheidet, und passt daraufhin die Gewichte an. Die Anpassung der Gewichte erfolgt so, dass der Unterschied zum gewünschten Ergebnis kleiner wird. Dieser Prozess wird wiederholt, bis die Fehler klein genug sind. Dabei werden nicht nur die Neuronen der Ausgabeebene angepasst, sondern die Informationen werden durch die Neuronenschichten zurückgeleitet, sodass jede Neuronenschicht lernt, ihre Ausgabe so zu adaptieren, dass der Fehler an der Ausgabe kleiner wird.

Die Anwendungen dieses Algorithmus blieben damals allerdings begrenzt, denn die Hardware war nicht leistungsfähig genug, um größere Netze zu verarbeiten, die komplexere Probleme bewältigen können. Es fehlten nicht nur Rechenressourcen, sondern auch die großen Datensätze, die wir heute selbstverständlich nutzen.

Mit den Fortschritten Mitte der 2000er-Jahre wurde deutlich, dass neuronale Netze mehr Potenzial hatten als ursprünglich gedacht. Das Ende des KI-Winters und die Wiederbelebung der Forschung an neuronalen Netzen begann 1986 mit den Arbeiten von David Rumelhart, Geoffrey Hinton und Ronald Williams. Sie betrachteten neuronale Netze mit »Hidden Layers«, also versteckten Schichten, die es den Netzen erlaubten, nichtlineare Probleme zu lösen, bei denen eine Veränderung der Eingabe nicht zu einer proportionalen Veränderung der Ausgabe führt. Diese versteckten Schichten waren der Schlüssel, um komplexere Muster in Daten zu erkennen und zu verarbeiten. Eine wichtige mathematische Erkenntnis dieser Zeit war, dass neuronale Netze mit genügend vielen Neuronen und

mindestens einer versteckten Schicht oder genügend vielen versteckten Schichten mit mindestens zwei Neuronen jede beliebige Funktion approximieren können.

In den frühen Tagen neuronaler Netze war es üblich, mit nur einer Hidden Layer zu arbeiten. Dies reichte für einfache Aufgaben aus, doch die begrenzte Tiefe schränkte die Fähigkeit des Netzes ein, komplexere Muster zu erkennen. Der große Durchbruch kam mit Netzen, die Dutzende oder sogar Hunderte von Schichten enthalten. Diese tiefen Netze, das sogenannte »Deep Learning«, konnten viel abstraktere Merkmale aus den Daten extrahieren, was ihre Anwendungsbereiche erheblich erweiterte.

In frühen Netzwerken dieser Art gab es nur wenige Dutzend Gewichte, die die Neuronen miteinander verbanden. Mit dem Aufkommen leistungsfähiger Hardware und umfangreicher Datenmengen konnten jedoch Netze mit Hunderten oder Tausenden von Gewichten trainiert werden.

Ein anschauliches Beispiel ist wieder die Klassifikation von Bildern mit Katzen und Hunden. Zu Beginn hat das Netz keinerlei Vorstellung von den relevanten Merkmalen, die es erlauben, Katzenbilder von Hundebildern zu unterscheiden. Es beginnt mit zufälligen Gewichten und passt diese Schritt für Schritt an, basierend auf den Fehlern, die es bei den Klassifikationen macht. Nach vielen Wiederholungen ist das Netz in der Lage, auch unbekannte Bilder korrekt zu klassifizieren, indem es Muster in den Pixeln erkennt.

Die automatische Generierung von Mustern führte zu großen Fortschritten in der Bild-, Text- und Sprachverarbeitung. In der Bildverarbeitung zeigten neuronale Netze, dass sie lernen können, komplexe Muster zu erkennen, ohne dass menschliches Eingreifen notwendig ist. Auch in der Sprach- und Textverarbeitung konnten sie komplexe Muster identifizieren, was maschinelle Übersetzungen, , Texterzeugung, Spracherkennung und andere Anwendungen verbesserte.

Vor der Ära der tiefen neuronalen Netze war das »Feature-Engineering« die vorherrschende Methode, d. h. Menschen haben sich »Features«, d. h. Muster bzw. Merkmale, ausgedacht. Expertinnen und Experten aus den

jeweiligen Fachgebieten entwickelten manuell Muster oder Merkmale, die für die jeweilige Aufgabe relevant sein könnten. In der Bildverarbeitung etwa wurde untersucht, welche Kanten, Formen oder Texturen als Muster oder Merkmale geeignet sein könnten, um Objekte zu erkennen. Diese Features wurden dann in Algorithmen integriert. Neuronale Netze revolutionierten diesen Ansatz. Statt Features manuell zu entwickeln, lernen sie diese automatisch aus den Daten. In den ersten Schichten eines Netzes werden grundlegende Merkmale wie Kanten erkannt. In den mittleren Schichten werden diese zu komplexeren Mustern kombiniert, etwa zu Linien, Halbkreisen oder Ovalen. Die obersten Schichten können schließlich spezifische Merkmale wie Augen, Schnauze oder Schnurrhaare einer Katze oder eines Hundes erkennen.

Diese Fortschritte stellten etablierte Methoden infrage und führten zu einem Paradigmenwechsel. Forschende, die zuvor Jahrzehnte in manuelles Feature-Engineering investiert hatten, mussten akzeptieren, dass neuronale Netze mit ausreichend Daten und Rechenleistung bessere Ergebnisse erzielen konnten – und das ohne tiefgehendes Domänenwissen.

Der Durchbruch von Deep Learning kam 2012, als Alex Krizhevsky, Ilya Sutskever und Geoffrey Hinton mit ihrem »AlexNet« den »ImageNet«-Wettbewerb gewannen, einen Wettbewerb, bei dem Forschende ihre Algorithmen zur Objekterkennung anhand einer umfangreichen, manuell annotierten Bilddatenbank testen. Ihr Deep Learning hatte alle Methoden, die von Expertinnen und Experten in der Bildverarbeitung entwickelt wurden, in den Schatten gestellt, ohne dass die Forschenden Wissen im Bereich der Bildverarbeitung hatten. Dieser AlexNet-Moment zeigte, dass neuronale Netze traditionelle Methoden der Bildverarbeitung deutlich übertreffen können – der Beginn einer neuen Ära der KI-Forschung.

Ein weiterer Meilenstein war die Einführung der Verwendung von GPUs (Graphics Processing Units) für das Training von neuronalen Netzen. Ursprünglich für die Grafikdarstellung in Computern entwickelt, erkannten Forschende bald, dass GPUs ideal für die parallele Verarbeitung von Daten geeignet waren. Insbesondere konnten große Matrixmultiplikationen

parallel gerechnet werden, wobei die Gewichte der Verbindungen in Matrizen gespeichert werden. Diese Fähigkeit der Parallelverarbeitung ermöglichte es, neuronale Netze mit Millionen von Parametern effizient zu trainieren. Plötzlich konnten Modelle entwickelt werden, die weitaus komplexer und leistungsfähiger waren als alles, was zuvor möglich gewesen war.

Zu leistungsfähigeren Computerchips kam die Verfügbarkeit großer Datensätze hinzu. In den frühen Jahren der KI war es oft schwierig, genügend Daten zu sammeln, um ein Netz so zu trainieren, dass es eine Aufgabe hinreichend gut lösen konnte. Doch mit dem Aufstieg des Internets und der Digitalisierung entstanden riesige Mengen an Daten, die neue Möglichkeiten eröffneten. Datensammelprojekte, wie das oben erwähnte ImageNet, wurden zu entscheidenden Katalysatoren für die Forschung.

Lernalgorithmen, insbesondere der bereits genannte Backpropagation-Algorithmus, wurden weiter optimiert und für große Netze angepasst. Neue Regularisierungstechniken halfen, das Problem des Overfittings zu bewältigen, bei dem ein Modell die Trainingsdaten auswendig lernt und dadurch bei neuen Daten schlecht generalisiert. Eine neue Regularisierungstechnik war damals »Dropout«, bei der im Netz während des Trainings zufällig Aktivierungen von Neuronen auf null gesetzt werden, wodurch nur robuste und häufig vorkommende Muster gelernt werden können und somit Auswendiglernen vermieden wird.

Für mich persönlich war diese Zeit besonders inspirierend. Ich konnte miterleben, wie sich die theoretischen Grundlagen, die in den Jahrzehnten zuvor gelegt worden waren, plötzlich in greifbare Anwendungen verwandelten – und wie die neuronalen Netze aus dem Schatten der symbolischen KI heraustraten.

Besonders beeindruckend war der Erfolg in der Bild-, Sprach- und Textverarbeitung. Neuronale Netze konnten nun Objekte in Bildern erkennen und korrekt klassifizieren sowie später sogar Bilder generieren. In der medizinischen Diagnostik zeigten neuronale Netze ihre Stärke. Sie konnten Muster in Röntgenbildern erkennen, die selbst erfahrene

Radiologen übersehen hätten, oder Hautkrebs besser als Dermatologen diagnostizieren. Das waren Erfolge, die weit über die reine Wissenschaft hinausgingen und die Grundlagen für zahlreiche kommerzielle Anwendungen schufen.

Ein weiterer Bereich, in dem neuronale Netze glänzten, war die Spracherkennung. Systeme wie Siri oder Alexa basieren auf Technologien wie Long Short-Term Memory (LSTM), die in den 1990er-Jahren entwickelt wurden. Doch erst mit der Einführung tiefer Netze und der Fähigkeit, größere Datenmengen zu verarbeiten, wurden diese Systeme gut genug, um sie zu nutzen.

Was mich besonders fasziniert, ist die Vielfalt der heutigen Anwendungen. Von der Bild- und Spracherkennung über Klimaforschung, Wettersimulationen oder Kernfusion bis hin zu selbstfahrenden Autos – all diese Anwendungen von KI-Technologien basieren auf Konzepten, die in den frühen Jahren der KI-Forschung entwickelt wurden.

Die Bedeutung neuronaler Netze für die heutige KI-Forschung und -Anwendung ist unbestreitbar. Doch ihre Entwicklung wirft auch grundlegende Fragen auf: Wie weit können wir diese Systeme noch verbessern? Welche neuen Paradigmen werden entstehen, wenn die Grenzen der aktuellen Ansätze erreicht sind?

2

KI mit Gedächtnis

Neuronale Netze hatten lange Zeit ein Gedächtnisproblem. Sie konnten nicht lernen, sich etwas aus der Vergangenheit zu merken, wenn es zu weit zurücklag. Nehmen Sie zum Beispiel den Satz: »Ein Mann im roten Pulli geht über die Straße.« Und jetzt denselben Satz in einer anderen Reihenfolge: »Über die Straße in einem roten Pulli geht ein Mann.« Die damaligen neuronalen Netze – insbesondere die »rekurrenten« Netze, die Informationen speichern können – hätten nur in der zweiten Variante erkannt, dass es sich um einen Mann handelt – also wenn das Wort am Ende der Sequenz steht. Länger zurückliegende Informationen konnten sie nicht abspeichern. Erst der von mir 1991 in meiner Diplomarbeit entwickelte Long Short-Term Memory (LSTM)-Algorithmus gab neuronalen Netzen ein funktionierendes Gedächtnis, das länger in der Zeit zurückreicht und sich auch Wörter vom Satzanfang merken kann.

Die Entwicklung von LSTM war keine rein theoretische Aufgabe. Tatsächlich begann alles mit praktischen Experimenten anhand von Sequenzen. Das rekurrente neuronale Netz verarbeitet die Elemente der Sequenz in Zeitschritten von vorne nach hinten. Ich hatte beobachtet, dass rekurrente Netze solche Sequenzen bei bestimmten Aufgaben nicht gut lernten. Ich wollte herausfinden, warum das so war. Als ich die Netze und ihr Verhalten genauer untersuchte, bemerkte ich, dass sie kaum in der Lage waren, Informationen aus den frühen Zeitschritten im Gedächtnis zu behalten. Der Grund dafür war das sogenannte »Vanishing-Gradient-Problem«. Beim Lernprozess durch Backpropagation wird berechnet, wie stark die Gewichte des Netzes angepasst werden müssen, um den Fehler zu verringern. Doch je weiter man in der Sequenz zurückgeht, desto schwächer

wird dieses Signal – der sogenannte »Gradient« –, bis es schließlich fast vollständig verschwindet. Dadurch gehen wichtige Informationen verloren, die eigentlich für das Verständnis der gesamten Sequenz entscheidend wären – wie im obigen Beispiel die Information, dass es sich um einen »Mann« handelt, der die Straße überquert.

Ich erkannte schnell, dass dieses Problem gelöst werden musste, um rekurrente Netze praktikabel zu machen. Mit meinem mathematischen Hintergrund konnte ich das Phänomen formal beschreiben und verstehen. Die Ursache lag in der wiederholten Multiplikation von Matrizen, die die synaptischen Verbindungen im Netz repräsentieren. Bei jeder Multiplikation wurden die verarbeiteten Werte weiter verkleinert, sodass am Ende die Information zu klein war, um zu lernen. Diese mathematische Einsicht war entscheidend, um eine Lösung zu entwickeln.

Die Herausforderung bestand nun darin, eine Architektur zu entwerfen, die das Vanishing-Gradient-Problem löste. Dabei half mir mein praktischer Zugang: Ich konnte die abstrakten mathematischen Konzepte mit konkreten Algorithmen und Daten in Verbindung bringen. Der Schlüssel war die Einführung von Speicherzellen, die Informationen über längere Zeiträume hinweg unverändert halten konnten. Speicherzellen haben Matrizen, die bei der Multiplikation die verarbeiteten Werte nicht verkleinern, das heißt, die Werte, die zu frühen Zeitschritten durch Multiplikation zurückgereicht werden, sind so groß wie die Werte von späten Zeitschritten. Da zurückpropagierte Fehler nicht verkleinert werden, nannten wir diesen Mechanismus »Constant Error Carousel«. Deshalb können Speicherzellen Informationen exakt über lange Zeiträume speichern. Zusätzlich benötigt man Mechanismen, um zu kontrollieren, welche Informationen in den Speicher gelangen und welche genutzt werden sollten. Diese Mechanismen nannten wir »Gating«. So entstanden die zwei zentralen Komponenten eines LSTM: »Constant Error Carousel« und »Gating«.

Das LSTM besitzt drei Gates. Das »Input Gate« entscheidet, welche Informationen in die Speicherzelle aufgenommen werden. Es verhindert, dass unwichtige Daten den Speicher überladen. Da alles exakt gespeichert

wird, wird eben auch alles Unwichtige gespeichert, was das Input Gate verhindert. Das »Output Gate« steuert, welche gespeicherten Informationen für die nächste Verarbeitungsschicht freigegeben werden. Später wurde von der Gruppe um Jürgen Schmidhuber noch das »Forget Gate« eingeführt, das festlegt, ob die gespeicherten Informationen vermutlich in der Zukunft weniger relevant sind und daher runterskaliert werden können.

Diese LSTM-Speicherzellen-Architektur ermöglicht es, wichtige Informationen über längere Zeiträume hinweg zu speichern, ohne dass sie durch neue Eingaben überschrieben oder vergessen werden. Gleichzeitig können unwichtige Daten gezielt vermieden werden, um unerwünschte Überlagerungen zu vermeiden. Ein neuronales Netz mit diesen Speicherzellen wurde »LSTM-Netz« oder kurz »LSTM« genannt.

Das LSTM revolutionierte die Analyse und Vorhersage basierend auf Zeitreihen und Sequenzen. Nun war es möglich, lange Sequenzen zu verarbeiten und wichtige Informationen über größere Zeiträume hinweg zu speichern. Anwendungen, die zuvor als unlösbar galten, wurden durch LSTM-Netze möglich. Besonders in der Textverarbeitung zeigten sich deren Stärken. Sie konnten außerdem den Kontext eines Satzes besser erfassen, was zu bedeutenden Fortschritten in der maschinellen Übersetzung, Spracherkennung und Textgenerierung führte.

LSTMs zeigten, dass rekurrente Netze nicht nur theoretisch, sondern auch praktisch leistungsfähig sind. Jedoch blieben sie viele Jahre unbeachtet. Erst um 2012 führten sie zu einem Paradigmenwechsel in der Sprach- und Textverarbeitung: Immer mehr Forscher und Entwickler setzten auf LSTMs und verwandte Architekturen.

Die Arbeit an LSTM führte mir vor Augen, wie wichtig die Kombination von Intuition und mathematischer Analyse ist. Praktische Experimente lieferten die notwendigen Beobachtungen, um die Probleme zu identifizieren, während die Mathematik half, diese zu formalisieren und dann systematisch zu lösen. Viele der Erkenntnisse wären ohne diese enge Verknüpfung nicht möglich gewesen. Ich habe oft beobachtet, dass manche Forschende zwar die mathematischen Fähigkeiten hatten, aber ohne

praktische Daten nicht wussten, wie sie ihre Fähigkeiten anwenden sollten. Andere hatten ein gutes Gespür für Daten, konnten aber ihre Beobachtungen nicht in mathematische Beschreibungen überführen. Der Erfolg von LSTM zeigte mir den entscheidenden Vorteil einer Verbindung beider Perspektiven.

Eine unsichtbare Revolution

Für mich persönlich war die Entwicklung von LSTM eine prägende Erfahrung. Sie führte mir vor Augen, wie wichtig es ist, Probleme ganzheitlich anzugehen – mit praktischen Experimenten, fundierter mathematischer Analyse und einem klaren Ziel vor Augen.

LSTMs erwiesen sich als revolutionär für viele Anwendungen, die wir heute täglich nutzen. Sie ermöglichten es, lange Sequenzen zu analysieren und Informationen über Zeiträume hinweg zu speichern. Besonders in der Sprachverarbeitung zeigten LSTMs ihre Stärken, wenn akustische Signale auf Texte abgebildet wurden. In der Textverarbeitung konnten sie den Kontext eines Satzes besser erfassen als bisherige Ansätze, was später zu bedeutenden Fortschritten bei maschineller Übersetzung, Textanalyse und Textgenerierung führte.

Doch die meisten Menschen wissen gar nicht, dass diese Technologie existiert. Dies spiegelt einen größeren Trend wider: Viele der wichtigsten Fortschritte in der KI geschehen im Verborgenen. Sie verbessern Anwendungen auf fundamentale Weise, bleiben aber für den Durchschnittsnutzer unsichtbar.

Ein zentraler Grund für die ursprüngliche Ablehnung neuronaler Netze waren ihre technischen Beschränkungen. Rechenleistung war nicht vorhanden, Datensätze waren klein und die Algorithmen waren nicht auf große Netze ausgelegt. Hinzu kam das schon erwähnte Problem des Vanishing Gradient, das dazu führte, dass Signale in tiefen Netzen verschwanden. Der Vanishing Gradient ist nicht nur, wie oben beschrieben, bei rekurrenten Netzen ein Problem, wo man über die Zeitschritte

propagiert, sondern auch bei tiefen Netzen, wo man über die Schichten propagiert.

1994 reichten wir das erste Manuskript zu LSTM bei der renommierten NeurIPS-Konferenz ein – es wurde abgelehnt. Auch war die Begeisterung für neuronale Netze Mitte der 1990er-Jahre abgeflaut. Viele Forschende waren enttäuscht von den begrenzten Möglichkeiten dieser Technologie und hatten sich neuen Ansätzen wie den »Support Vector Machines« zugewandt. Die wissenschaftliche Gemeinschaft nahm mein Konzept zunächst nicht wahr. Ich war jung und dachte mir, dass andere Ansätze vielleicht tatsächlich besser wären. Obwohl ich wusste, dass LSTM funktionierte, konnte ich damals die Tragweite dieser Technologie nicht einschätzen. Ich erinnere mich an einen Workshop bei der NeurIPS-Konferenz, bei dem ich versuchte, die Ergebnisse meiner LSTM-Forschung zu präsentieren. Die Reaktion war ablehnend: »Neuronale Netze? Die funktionieren nicht. Wir wollen neue Technologien erforschen.« Neuronale Netze galten als veraltet. Es war frustrierend, aber ich blieb davon überzeugt, dass meine Arbeit mit LSTM gute Ergebnisse lieferte.

Um mich in der Forschung weiterzuentwickeln, wandte ich mich ebenfalls den damals populären Forschungsansätzen wie den genannten »Support Vector Machines« zu. Doch hier war ich spät dran: Andere Forschende hatten in diesen Bereichen oft Jahre Vorsprung. Ich war in der Machine-Learning-Community zu einem Außenseiter geworden, weil mein Thema als altmodisch galt. Bei den neuen Trends hingegen war ich nicht früh genug eingestiegen. In der Bioinformatik fand ich eine Nische, in der ich meine Machine-Learning-Methoden einsetzen konnte – sowohl die veralteten neuronalen Netze wie LSTM als auch damals moderne Methoden wie Support Vector Machines und »Kernel Methods«. 2006 erhielt ich eine Professur in Linz, die mir große Freiheiten bot. Ohne den Druck, für eine Uniklinik oder andere Institutionen spezifische Leistungen bereitzustellen, konnte ich mich meiner Forschung widmen – sei es in der Bioinformatik oder im »Machine Learning«.

Ein Kapitel meiner Karriere, das mir besonders wichtig ist, war meine Rückkehr zum maschinellen Lernen nach einigen Jahren in der Bioinformatik. In den frühen 2000er-Jahren hatte ich mich auf biologische Anwendungen konzentriert, weil ich dort spannende Herausforderungen sah, die durch maschinelles Lernen gelöst werden konnten. Doch die Einführung von Deep Learning und die Wiederentdeckung neuronaler Netze hatten das Feld von Machine Learning revolutioniert. Es war klar, dass sich in diesem Feld etwas Grundlegendes veränderte, man sprach nun von »künstlicher Intelligenz« oder »KI«. Das war für mich der Moment, in dem ich entschied, wieder stärker in die Machine-Learning-Forschung, nun die KI-Forschung, einzusteigen.

Während ich in Linz arbeitete, begann sich das Feld der neuronalen Netze langsam zu verändern. Ab 2010 wurden tiefere Netze und größere Datenmengen plötzlich relevant. LSTMs, die lange Zeit wenig Beachtung gefunden hatten, rückten überraschend ins Zentrum der Aufmerksamkeit. Ich erinnere mich, wie ich bei einer NeurIPS-Konferenz feststellte, dass ich nun sogar sehr bekannt war. Geoffrey Hinton, Turing-Preisträger 2019 und Nobelpreisträger für Physik 2024, kam auf mich zugelaufen und sagte: »Sepp, wegen dir sind nun all diese Leute auf der NeurIPS.« Früher hatte ich immer Herzklopfen, wenn Hinton vorbeiging, und ich hätte mir nie erträumen können, dass er mich einmal ansprechen würde. Auf dieser NeurIPS wollten plötzlich Leute Selfies mit mir machen und fragten mich, wie ich auf LSTM gekommen war. Ich hatte zunächst nicht realisiert, dass meine frühen Beiträge nun als revolutionär wahrgenommen wurden.

Es war ein überwältigendes Gefühl, als die wissenschaftliche Gemeinschaft LSTM schließlich als entscheidende Erfindung anerkannte. Technologien wie Handschrifterkennung oder Sprachverarbeitung bauten darauf auf, und plötzlich war LSTM ein unverzichtbarer Bestandteil moderner KI-Anwendungen.

Der endgültige Durchbruch für LSTM kam, als große Tech-Unternehmen wie Google, Apple und Amazon begannen, die Technologie in ihren Produkten zu verwenden. Zwischen 2010 und 2017 dominierte LSTM viele

Bereiche der KI. Google galt in dieser Zeit als ein »LSTM-Unternehmen«, da viele seiner Anwendungen auf dieser Technologie basierten – insbesondere in der Text- und Sprachverarbeitung. Auch einige der Gründer von OpenAI wie Ilya Sutskever, Wojciech Zaremba und Andrej Karpathy, die Forscher hinter ChatGPT, arbeiteten damals – vor dem Durchbruch der Transformer-Architektur – vor allem mit LSTM. Auf der NeurIPS im Dezember 2015 sagten Ilya Sutskever, Wojciech Zaremba und Andrej Karpathy zu mir: »Du musst kommen, wir haben eine wichtige Ankündigung zu machen.« Dabei handelte es sich um die Bekanntgabe der Gründung von OpenAI.

LSTM ist für verschiedenste Anwendungen geeignet: von »Predictive Maintenance«, also der proaktiven Instandhaltung von Maschinen und Anlagen, über Finanzprognosen bis hin zur Spracherkennung. Hier waren LSTM-Netze lange Zeit unschlagbar.

Zwei Mojitos von Amazon

Eine Frage, die mir oft gestellt wird, lautet: »Haben Sie jemals von LSTM profitiert?« Die kurze Antwort ist: nicht finanziell. Als ich die Technologie entwickelte, war ich ein junger Student, und die Idee, ein Patent anzumelden, kam mir nicht in den Sinn – vor allem, weil LSTM zunächst auf Ablehnung stieß. An eine kommerzielle Verwertung war damals nicht zu denken. Als ich LSTM entwickelte, kämpfte ich darum, dass die Technologie überhaupt in der wissenschaftlichen Gemeinschaft akzeptiert wurde. Ich kannte nur den akademischen Ansatz: Ergebnisse publizieren und das Forschungsfeld voranbringen. Später bauten große Unternehmen wie Google plus DeepMind, Meta, Microsoft, Apple, Amazon, Baidu, Tencent, Samsung, Huawei usw. auf den Grundlagen meiner Forschung auf und entwickelten spezialisierte Architekturen. Ich dachte nie daran, eine Firma zu gründen. Wichtiger war mir die Freude, an der Universität neue

Ideen zu entwickeln. Wissenschaftliche Probleme waren wie große Rätsel für mich – Rätsel, die ich unbedingt lösen wollte.

Ich bereue die Entscheidung nicht, dass ich eine akademische Laufbahn eingeschlagen habe. Wäre LSTM patentiert worden, um damit eventuell Geld zu machen, hätten viele Forschende und Unternehmen möglicherweise gezögert, die Technologie zu übernehmen. LSTM hätte nie einen Durchbruch gehabt. Stattdessen wurde LSTM ein integraler Bestandteil der KI-Community, und ich konnte von der Anerkennung meiner Arbeit profitieren, die mir eine Plattform für weitere Forschungen bot. Der Erfolg von LSTM führte dazu, dass immer mehr Forschende sich mit diesem Ansatz beschäftigten. Ich war zwar nicht immer aktiv in alle Entwicklungen eingebunden, aber ich unterstützte die Forschenden, die auf meinen Grundlagen aufbauten.

Manchmal haben wissenschaftliche Entdeckungen überraschende Folgen. Ich erinnere mich an eine Begegnung mit Amazon-Mitarbeitern im Jahr 2018. Sie berichteten mir, dass ihre Implementierung einer von Günter Klambauer und mir entwickelten Methode – der Self-Normalizing Networks (SNN) – ihren Umsatz um eine Milliarde Dollar gesteigert hatte. Als Dankeschön luden sie uns auf jeweils einen Mojito ein. Amerika: eine Milliarde Dollar Umsatzsteigerung; Europa: zwei Mojitos.

Ähnlich war es mit Apple: Dort wurde dieselbe Technologie verwendet, um neuronale Netze so zu optimieren, dass sie auf iPhones verwendet werden konnten. Wir bekamen eine E-Mail vom iOS-Chefentwickler, der sich für die Erfindung von SNNs bedankte, da sie Apple so viel geholfen hatte. Diese Beispiele verdeutlichen, wie kleine Änderungen in der Forschung massive Auswirkungen in der Industrie haben können. Für mich war es immer eine Freude, zu sehen, wie meine Arbeiten genutzt wurden – auch wenn ich davon nicht finanziell profitierte.

Die Geschichte von LSTM zeigt, wie wichtig es ist, an wissenschaftlichen Ideen festzuhalten, auch wenn sie zunächst nicht gewürdigt werden. Die späte Anerkennung meiner Arbeit hat mir gezeigt, dass wahre Innovation oft Zeit braucht, um sich durchzusetzen. Heute bin

ich stolz darauf, einen Beitrag zur Entwicklung moderner KI geleistet zu haben, der die Welt verändert hat.

Siegeszug der Transformer

Trotz ihres Erfolgs wurden LSTM-Netze schließlich von Transformer-Modellen wie GPT (Generative Pretrained Transformer) und BERT (Bidirectional Encoder Representations from Transformers) abgelöst. Der entscheidende Vorteil der Transformer war, dass sie an großen Datenmengen effizienter trainiert werden konnten. Insbesondere der neue Mechanismus der »Attention« wurde in Transformern durch Parallelisierung optimiert, damit waren Transformer-Architekturen im Training schneller als LSTM-Architekturen. Der Attention-Mechanismus erlaubt neuronalen Netzen, sich gezielt auf die relevantesten Teile einer Eingabesequenz zu fokussieren, anstatt alle Elemente gleich zu gewichten. So erfasst der Self-Attention-Mechanismus in Transformern den Zusammenhang zwischen einem weit in der Vergangenheit liegenden Wort und dem aktuellen Wort in einem Text. Da jedes vergangene Wort erst exakt gespeichert und dann direkt betrachtet wird, gibt es kein Vanishing-Gradient-Problem.

Das Paper *Attention is All You Need* im Jahr 2017 markierte den Beginn einer neuen Ära. Transformer-Modelle waren besser parallelisierbar. Das bedeutet, dass Trainingsberechnungen für ein Modell gleichzeitig auf vielen Prozessoren oder Recheneinheiten ausgeführt werden können. Dies ist ein entscheidender Vorteil, insbesondere in der heutigen Zeit, in der große Datenmengen (Big Data) verarbeitet werden müssen. Dennoch blieb LSTM in spezifischen Anwendungen, wie bei Reinforcement Learning (AlphaStar oder OpenAI 5), der Sprachverarbeitung und Sprachsynthese, der Analyse von Finanzdaten oder allgemein bei Zeitreihen, bis heute die beste Methode.

Transformer wurden schnell zur dominierenden Technologie in der KI-Forschung, insbesondere in den Bereichen Textverstehen und -generierung.

Ihre Fähigkeit zur Skalierung und ihre Effizienz machten sie zur bevorzugten Wahl in vielen Anwendungen. Die Aufmerksamkeit der KI-Gemeinschaft lag auf dem Neuen, wohingegen ältere Ansätze wenig beachtet wurden. Die Wissenschaft ist oft von Trends und Moden geprägt. Wenn eine neue Methode aufkommt, richtet sich die Aufmerksamkeit der Gemeinschaft auf diese. Das war bei den Transformern nicht anders. Die Geschwindigkeit, mit der sich die Forschungsgemeinschaft auf diese Technologie stürzte, war beeindruckend – und auch ein wenig entmutigend für diejenigen, die an alternativen Ansätzen arbeiteten.

Die Ablösung von LSTM durch Transformer hat mir gezeigt, wie wichtig Timing in der Wissenschaft ist. Es reicht nicht, eine gute Idee zu haben; sie muss auch zur richtigen Zeit kommen. LSTM hatte klare Vorteile, insbesondere bei der Modellierung von Zeitreihen. Doch ihre Skalierbarkeit zu größeren Modellen und mehr Daten war begrenzt, und das machte sie in einem Zeitalter, in dem Rechenleistung und Daten immer wichtiger wurden, weniger attraktiv. Transformer sind effizient im Training und gut auf große Modelle skalierbar, weshalb sie sich durchsetzten.

Doch Timing allein reicht nicht aus. Es ist ebenso wichtig, flexibel zu bleiben und sich an neue Entwicklungen anzupassen. Die KI-Forschung ist ein schnelllebiges Feld, und wer nicht bereit ist, neue Ansätze zu erforschen, riskiert, den Anschluss zu verlieren. Ich frage mich oft: Hätten wir LSTM-Netze früher skalieren können? Hätten wir sie besser präsentieren müssen? Diese Fragen sind schwer zu beantworten. Doch sie zeigen, wie wichtig es ist, ständig an der Verbesserung einer Technologie zu arbeiten – auch wenn sie bereits etabliert ist.

Die Einführung der Transformer war ein Wendepunkt in der KI. Als die ersten Arbeiten zu dieser Technologie veröffentlicht wurden, erkannte ich nicht sofort ihr volles Potenzial. Es war eine graduelle Entwicklung: Zunächst waren Transformer einfach eine weitere Methode, doch nach und nach zeigten sie ihre Überlegenheit in bestimmten Bereichen, insbesondere bei Texten.

Transformer wurden schnell zu einem Hype, und viele Forscher wechselten zu dieser neuen Technologie. Gleichzeitig bedeutete das, dass andere Ansätze – selbst solche mit großem Potenzial – weniger Beachtung fanden. Ich erinnere mich an Gespräche mit Kollegen, die sagten: »Ich bekomme keine Finanzierung, wenn ich an älteren Technologien arbeite. Das Geld fließt in die neuen Ansätze.« Dieses Phänomen ist typisch für die Wissenschaft: Sobald ein Ansatz als »der nächste große Wurf« wahrgenommen wird, richten sich alle Augen auf ihn. Das ist nicht unbedingt schlecht. Hypes können den Fortschritt beschleunigen, weil sie Ressourcen und Aufmerksamkeit auf ein bestimmtes Thema lenken. Doch sie können auch dazu führen, dass andere wichtige Ideen übersehen werden.

Teamarbeit – der Schlüssel zur Innovation

Wissenschaft ist nie eine Einzelleistung. Jeder bringt seine Stärken und Perspektiven ein, und gemeinsam können wir Probleme lösen, die für einen Einzelnen unüberwindbar wären. Es gibt zwar Momente, in denen ich mich allein in mein »Kämmerchen« zurückziehe, um Analysen durchzuführen oder neue Ideen zu entwickeln. Doch der Großteil der Arbeit, insbesondere in der modernen KI, findet in Teams statt.

In meinen Teams haben wir oft mit verschiedenen Architekturen experimentiert. Manchmal waren es Methoden von anderen Forschungslabs, die uns auf eine neue Idee brachten. Diese Dynamik zeigt, wie wichtig es ist, offen für neue Ansätze zu sein und voneinander zu lernen. In meiner Karriere hatte ich das Glück, mit vielen talentierten Forschern zusammenzuarbeiten. Was mir besonders am Herzen liegt, ist die Förderung junger Wissenschaftler. Ich sehe es als meine Aufgabe, ihnen Raum zu geben, ihre eigenen Ideen zu entwickeln und zu verfolgen. Das bedeutet auch, dass ich mich manchmal bewusst zurücknehme. Ich bin nicht auf jedem Paper als Co-Autor vertreten, selbst wenn ich an der Arbeit beteiligt

war. Ich möchte, dass meine Kollegen ihre eigenen Erfolge feiern können, ohne dass mein Name ihre Leistung überschattet.

Forschung ist nicht nur rational. Monatelanges Arbeiten an einem Problem bringt auch emotionale Höhen und Tiefen mit sich. Erfolge und Rückschläge prägen den Prozess und treiben uns an. Erfahrung ist in der Wissenschaft unbezahlbar und führt schließlich zu einer Intuition, die Theorie und Praxis verbindet. Ich erinnere mich an eine Situation, in der mir ein Mitarbeiter seine Ergebnisse präsentierte. Er zeigte die erste Folie mit einer Lernkurve, und ich konnte sofort erkennen, was er später in seinem Vortrag ausführlich erläutern wollte. Die Kurve zeigte, dass das Modell zwar schneller lernte als andere, aber sich zu sehr auf einfache Muster fokussierte und später die komplexen Muster nicht mehr lernte. Eine solche Intuition entsteht durch unzählige Stunden der Analyse und Tests von neuronalen Netzen. Jede Kurve erzählt eine Geschichte, die man lesen lernen muss. Manchmal probiert man eine Idee aus und das Modell liefert Ergebnisse, die nicht zu den Erwartungen passen. Hier wird es spannend, und man beginnt, zu überlegen und Hypothesen aufzustellen.

Die Wissenschaft ist ein nie endender Prozess des Lernens und Entdeckens. Sie erfordert Neugier, Durchhaltevermögen – und die Fähigkeit, aus Fehlern zu lernen. Ein weiterer faszinierender Aspekt ist die Rolle des Zufalls. Oft sind es unerwartete Ergebnisse, die uns auf neue Ideen bringen. Diese Zufälle sind jedoch nur dann wertvoll, wenn wir sie erkennen und daraus lernen.

Meine Zusammenarbeit mit Jürgen Schmidhuber ist ein wichtiger Teil meiner wissenschaftlichen Laufbahn. Schmidhuber war zwar nicht mein Doktorvater, jedoch spielte er als mein Betreuer die Hauptrolle während meiner Promotionszeit. Er hatte viele Ideen und war ein begeisterter Forscher. Nachdem Schmidhuber nach Lugano gegangen war, rief er mich an, um zu sehen, was bei LSTM noch fehlte oder wo es noch Probleme gab. Das war die Geburtsstunde der Idee des bereits erwähnten »Forget Gates«, das dann von Felix Gers und Jürgen Schmidhuber entwickelt wurde. Das

Forget Gate sollte dem LSTM erlauben, irrelevante Informationen gezielt zu vergessen.

Schmidhuber war eine prägende Figur, doch es war mir von Anfang an klar, dass ich meinen eigenen Weg gehen musste. Später betrieb ich von Schmidhuber unabhängige Forschung. Dennoch blieben wir immer in Kontakt und hatten gegenseitig großen Respekt voreinander. Auch wenn wir unterschiedliche Wege einschlugen, war mir unser Austausch immer wichtig.

Die Zukunft der KI wird von der Integration verschiedener Ansätze geprägt sein. Es wird darum gehen, die Stärken verschiedener Methoden zu vereinen und dabei über den Tellerrand der aktuellen Trends hinauszublicken. Diese Perspektive erfordert Mut, Kreativität und die Bereitschaft, bestehende Paradigmen infrage zu stellen. Die Wissenschaft ist ein ständiger Balanceakt zwischen Tradition und Innovation. Es geht darum, neue Ideen zu entwickeln, ohne die Errungenschaften der Vergangenheit zu vergessen, Trends zu hinterfragen und sich nicht von ihnen dominieren zu lassen.

Forschung ist für mich eine intellektuelle Herausforderung. Es geht darum, komplexe Probleme zu lösen, neue Ideen zu entwickeln und bestehende Technologien zu verbessern. Dies führt zu einem Wettbewerb unter den Forschungsgruppen: Wer löst die Probleme als Erstes und wie gut? Werden sie mit den eigenen Ansätzen gelöst?

Ein Beispiel ist meine Arbeit mit Transformern. Statt mich dagegen zu wehren, dass sie LSTM ersetzten, wollte ich verstehen, warum sie so erfolgreich waren. Diese Analyse führte zu neuen Erkenntnissen.

Ein zentraler Bestandteil der Forschung ist das Wechselspiel zwischen Theorie und Praxis. Oft beginnen Forschungsprojekte mit einer theoretischen Überlegung, die dann in der Praxis getestet wird. Dabei zeigt sich, ob die Theorie Bestand hat – oder angepasst werden muss. Andererseits gibt es Probleme in der praktischen Anwendung, wofür dann mathematische Ansätze gefunden werden, die wiederum zu theoretischen

Lösungen führen, die am Ende neue Algorithmen in der Praxis ermöglichen.

Ein Beispiel ist die Analyse von Attention-Mechanismen in Transformern. Zuerst wurde der Attention-Mechanismus in LSTM verwendet. Dann sah man, dass man das LSTM nicht mehr brauchte und Attention ausreichend war. In der Praxis zeigte sich, dass Attention vor allem die Trainingseffizienz und Skalierbarkeit der Modelle verbesserte. Diese Erkenntnis war entscheidend für die Weiterentwicklung von KI-Technologien.

Ich bin optimistisch, dass wir in den nächsten Jahren bedeutende Fortschritte sehen werden – nicht nur durch neue Technologien, sondern auch durch die Wiederentdeckung und Weiterentwicklung älterer Ansätze. Die Zukunft der KI gehört denen, die bereit sind, aus der Vergangenheit zu lernen und mutig nach vorne zu schauen.

Mehr Offenheit

Eine der erfreulichen Entwicklungen in der KI ist die zunehmende Verbreitung von Open-Source-Projekten. Diese Initiativen ermöglichen es Forschern und Unternehmen, auf vorhandenen Modellen und Software aufzubauen, anstatt jedes Mal von Grund auf neu zu beginnen.

Viele Wissenschaftler, darunter auch ich, sind der Ansicht, dass offene Modelle mit offenen Datensätzen, mit denen sie trainiert wurden, entscheidend sind, um die Forschung voranzutreiben. Aber auch die Trainings- und Inferenzsoftware sollte Open Source sein, um die Modelle nachbauen zu können. Open Source ermöglicht Transparenz: Jeder kann nachvollziehen, wie ein Modell trainiert wurde, welche Daten verwendet wurden und welche Mechanismen zugrunde liegen. Mit Open Source wird Forschung aufbauend auf den bekannten Modellen erst ermöglicht. Ein Merkmal der wissenschaftlichen Erkenntnis ist die Reproduzierbarkeit, die eben Open Source verlangt.

Transparenz ist auch aus ethischer und sicherheitstechnischer Perspektive wichtig. Wenn beispielsweise religiöse oder politische Texte in Trainingsdaten in überdurchschnittlichen Proportionen enthalten sind oder gewisse Texte und Fakten gelöscht werden, kann dies zu unerwünschten Verzerrungen führen. Ein Open-Source-Ansatz, bei dem die Trainingsdaten und ihre Verwendung offengelegt wurden, ermöglicht es, diese Risiken zu erkennen und zu minimieren. Außerdem wird es erheblich einfacher, Modelle anzupassen und ihre Schwachstellen zu beseitigen.

Im Gegensatz dazu stehen proprietäre, also firmeneigene Modelle großer Technologieunternehmen, die oft in einer Black-Box-Form präsentiert werden. Ohne Einblick in die zugrunde liegenden Daten und Trainingsmechanismen bleibt unklar, ob die Modelle potenzielle Gefahren bergen. Typischerweise sind in den Trainingsdaten einer KI kulturelle oder nationale Besonderheiten oder Vorlieben vorhanden, die sich unbemerkt auf die Ausgaben des Modells auswirken. Ein Textmodell, das primär mit amerikanischen Texten trainiert wurde, wird Antworten geben, die stärker von amerikanischen Werten geprägt sind.

Die Debatte um proprietäre und offene Modelle zeigt, wie wichtig es ist, eine Balance zwischen Innovation und Verantwortlichkeit zu finden. Open Source erlaubt es, die Vor- und Nachteile sowie Gefahren von Modellen und Methoden besser zu erkennen. Offene Modelle fördern die Demokratisierung der KI und sorgen dafür, dass die Technologie nicht nur wenigen großen Akteuren vorbehalten bleibt.

3

Mensch versus Maschine

Künstliche neuronale Netze werden oft mit dem menschlichen Gehirn verglichen. Auf den ersten Blick ist das auch naheliegend: Beide arbeiten mit Netzwerken aus miteinander verbundenen »Knoten«, die Informationen verarbeiten und weiterleiten. Und tatsächlich ließen sich die ersten Forschenden, die neuronale Netze entwickelten und sie auch so nannten, von der Struktur biologischer Gehirne inspirieren. Die Idee war, ein System zu bauen, das wie ein Mensch lernen kann. Doch dieser Vergleich hat seine Grenzen. Denn trotz der äußeren Ähnlichkeiten unterscheiden sich künstliche neuronale Netze und das Gehirn in ihrer Funktionsweise und Art zu lernen. Die genaue Funktionsweise des Gehirns ist noch eine offene Forschungsfrage, die im Feld Computational Neuroscience beantwortet werden soll.

Ein Beispiel verdeutlicht diesen Unterschied: Während ein neuronales Netz enorme Datenmengen benötigt, um Aufgaben zu lernen und zu generalisieren, genügen dem menschlichen Gehirn oft wenige Beispiele, um Generalisierungen vorzunehmen. Ein Kind kann nach wenigen Beispielen verstehen, wie ein neues Kartenspiel funktioniert, während ein neuronales Netz Hunderte oder Tausende von Beispielen braucht, um ein ähnliches Verständnis zu entwickeln. Diese Fähigkeit zur schnellen Generalisierung aus wenigen Beispielen ist etwas, das unserer heutigen KI noch fehlt.

Gleichzeitig hat das Gehirn eine bemerkenswerte Flexibilität. Es kann zwischen verschiedenen Aufgaben wechseln, sich an neue Situationen anpassen und komplexe Entscheidungen treffen. Neuronale Netze hingegen sind in der Regel hochspezialisiert auf eine Aufgabenstellung. Ein System, das für die Erkennung von Gesichtern trainiert wurde, ist nicht in der Lage,

plötzlich eine Textübersetzung zu leisten. Jedes neue Szenario erfordert ein neues Netz oder zumindest ein umfangreiches Umtrainieren.

Trotzdem ist es nicht sinnvoll, das Gehirn als Modell für künstliche Intelligenz zu idealisieren. Biologische Systeme sind oft ungenau, langsam und haben vor allem eine beschränkte Kapazität. Künstliche neuronale Netze hingegen folgen mathematischen Prinzipien, die auf Optimierung und Leistungsfähigkeit ausgelegt sind. Sie sind nicht an die Einschränkungen eines biologischen Organismus gebunden – etwa Energieverbrauch oder physikalischen bzw. chemischen Grenzen – und können deshalb Aufgaben lösen, die das Gehirn in dieser Form nicht bewältigen könnte.

Der Vergleich zwischen neuronalen Netzen und dem menschlichen Gehirn wirft also eine spannende Frage auf: Was können wir von biologischen Systemen lernen, und wo müssen wir neue Wege gehen? Obwohl künstliche neuronale Netze das Gehirn als Vorbild hatten, haben sich die modernen KI-Methoden fast nichts vom Gehirn abgeschaut und sich unabhängig entwickelt.

Das menschliche Gehirn und neuronale Netze haben nicht nur unterschiedliche Ursprünge, sondern auch stark divergierende Funktionsprinzipien. Man weiß, dass das Gehirn nicht mit Backpropagation lernt – also Fehler nicht rückwärts durch das Netz weiterleitet, um die Stärke neuronaler Verbindungen anzupassen –, da die gemessene Signalausbreitung im Gehirn das nicht erlaubt. Andererseits ist Backpropagation die Methode, um große neuronale Netze zu trainieren, und das mit größter Präzision. Diese Unterschiede sind entscheidend, um die jeweilige Leistungsfähigkeit und die Grenzen der beiden Systeme zu verstehen.

Effizienzwunder Gehirn

Das Gehirn ist ein biologisches Wunderwerk, das mit etwa 20 Watt Energie auskommt – vergleichbar mit einer kleinen Glühbirne. Damit koordiniert es

nicht nur unsere Gedanken und Emotionen, sondern steuert auch den gesamten Körper. Neuronale Netze hingegen erfordern enorme Rechenressourcen und Energie. Die Trainingsphase eines modernen KI-Systems kann den Energieverbrauch einer Kleinstadt übersteigen. Der Vergleich macht deutlich, wie unterschiedlich die beiden Systeme aufgebaut sind: Während das Gehirn darauf ausgelegt ist, mit minimalen Ressourcen maximale Leistungen zu erreichen, sind neuronale Netze durch viele Daten und hohe Rechenleistung getrieben.

Während neuronale Netze durch ihre Optimierung auf spezifische Aufgaben brillieren, ist die Anpassungsfähigkeit des Gehirns unübertroffen. Menschen lernen, indem sie Erfahrungen verknüpfen und auf neue Kontexte anwenden. Ein Kind, das einmal ein Ballspiel gelernt hat, wird die grundlegenden Prinzipien leicht auf andere Sportarten übertragen können. Neuronale Netze müssen dagegen für jede neue Aufgabe von Grund auf trainiert werden. Es gibt jedoch auch Anpassungsansätze für neuronale Netze wie Transferlernen, Multitasklernen, Anpassung an Domainshifts etc. Es wurden Agenten vorgestellt, die viele verschiedene Aufgaben lösen, wie GATO »A Generalist Agent« von DeepMind. Was die Anpassungsfähigkeit an neue Aufgaben und Weltverständnis anbelangt, kommen diese Ansätze jedoch nicht an das menschliche Gehirn heran.

Auch LSTM speichert wie der Mensch nur die relevanten Informationen. Das »Input Gate« bestimmt, welche Information relevant ist und gespeichert werden soll und welche nicht. Dieser Ansatz ist enorm ressourceneffizient, da er die benötigte Rechenleistung erheblich reduziert.

Doch wie entscheidet das Gehirn, welche Informationen wichtig sind? Neurowissenschaftler gehen davon aus, dass Emotionen eine Schlüsselrolle spielen. Situationen, die starke Gefühle hervorrufen, werden mit höherer Wahrscheinlichkeit gespeichert. Zwar besitzt das LSTM ein »Input Gate«, das oft als Analogon zum »Dopamine Gating« im menschlichen Gehirn genannt wird, jedoch fehlen Analogons zu Emotionen, die das Gating steuern.

Ein weiterer entscheidender Unterschied ist die Art, wie beide Systeme mit Ungenauigkeiten, fehlender Information und Ungewissheit umgehen. Das menschliche Gehirn ist erstaunlich tolerant gegenüber fehlender Information. Es kann Lücken durch kreative Rekonstruktion oder logisches Denken überbrücken. Neuronale Netze hingegen sind fehleranfällig, wenn Informationen oder Sensordaten fehlen.

Wie funktioniert Lernen?

Lernen ist die Fähigkeit eines Systems, aus vergangenen Erfahrungen Wissen zu extrahieren und dieses auf neue, zukünftige Situationen zu übertragen.

Abstraktion – also die Fertigkeit, das Wesentliche beizubehalten und das Unwesentliche zu ignorieren – ist eine der wichtigsten Eigenschaften der menschlichen Intelligenz und unserer Lernfähigkeit. Diese beiden Komponenten ermöglichen es uns, komplexe Probleme zu lösen und in unbekannten Situationen fundierte Entscheidungen zu treffen. Gut zu abstrahieren ist für KI-Systeme jedoch nach wie vor eine große Herausforderung.

Der Mensch ist meisterhaft darin, aus spezifischen Erfahrungen allgemeine Prinzipien abzuleiten, also zu abstrahieren. Dadurch fallen viele Objekte oder Situationen aus der realen Welt in die gleiche abstrakte Kategorie, wie zum Beispiel ein Tisch: Ein Kind lernt, dass ein Tisch vier Beine und eine Platte hat, auf die man etwas stellen kann. Hat das Kind das einmal gelernt, kann es eine Vielzahl unterschiedlicher Tische erkennen – selbst wenn diese aus verschiedenen Materialien bestehen oder ungewöhnliche Formen haben.

Die Fähigkeit zur Abstraktion ist essenziell, um neue Informationen effizient einzuordnen und anzuwenden – insbesondere wenn es darum geht, komplexe Probleme zu lösen. Denn die Lösungsfindung ist viel effizienter, wenn man unwichtige Details weglässt, die nicht zur Lösung beitragen.

Neuronale Netze benötigen immense Datenmengen, um einfache Abstraktionsleistungen zu erbringen. KI-Systeme müssen Millionen von Datenpunkten verarbeiten, um abstrakte Muster zu lernen.

Die menschliche Fähigkeit zur Abstraktion hat tiefe evolutionäre Wurzeln. Insbesondere ist sie die Voraussetzung für Sprache. Wörter sind oft Abstraktionen der echten Welt, wie Tisch, Baum, Hund oder Katze. Unsere Abstraktionsfähigkeit ermöglicht es uns, Werkzeuge zu entwickeln, Strategien zu planen und komplexe soziale Strukturen aufzubauen. Planen wird erst möglich, wenn man abstrahiert. Dabei wird nicht jeder kleine Schritt, den man macht, geplant, sondern es werden abstrakte Teilziele gesetzt: Erst gehe ich zur U-Bahn, dann steige ich bei der Endstation aus, und dann nehme ich den Bus nach Hause. KI-Systeme haben nicht nur Schwierigkeiten, zu abstrahieren und zu planen, es fehlt ihnen vor allem das Weltverständnis, also das Verständnis, wie die soziale oder physikalische Welt funktioniert.

Um KI-Systeme leistungsfähiger zu machen, müssen neue Ansätze entwickelt werden, die Abstraktion, Planung und Weltverständnis näher an die menschlichen Fähigkeiten heranführen. Dies könnte durch hybride Systeme erreicht werden, die subsymbolische Methoden wie neuronale Netze mit symbolischen Methoden wie Logiksystemen oder Wissensrepräsentation kombinieren. Eine solche Entwicklung würde nicht nur die Effizienz von KI-Systemen steigern, sondern auch deren Zuverlässigkeit in komplexen, dynamischen Umgebungen erheblich erhöhen.

Die Fähigkeit, Pläne zu schmieden und hypothetische Szenarien zu entwerfen, ist eine zentrale Eigenschaft menschlicher Intelligenz. Sie erlaubt es uns, über die Gegenwart hinauszublicken und Entscheidungen zu treffen, die auf zukünftigen Möglichkeiten und Risiken basieren. Wir Menschen sind in der Lage, mentale Modelle der Welt zu entwickeln, die es uns ermöglichen, auch ohne direkte Erfahrung Vorhersagen zu treffen. Ein einfaches Alltagsbeispiel ist das Planen einer Reise: Wir können uns

vorstellen, welche Kleidungsstücke in einer bestimmten Klimazone sinnvoll sind, auch wenn wir noch nie dort gewesen sind.

KI-Systeme haben Schwierigkeiten, längere Abläufe zu planen. Selbst fortgeschrittene Modelle wie GPT-4 oder DeepSeek R1 können zwar Texte generieren, aber ihre Planungsfähigkeit beschränkt sich auf Planungsbeispiele, die in den Trainingsdaten enthalten sind. Die Integration von Planungsfähigkeiten in KI erfordert Fortschritte in mehreren Bereichen. Ein Ansatz könnte die Kombination von Reinforcement Learning mit symbolischen Methoden sein, um langfristige Strategien zu entwickeln. Ein weiterer Ansatz ist die Entwicklung von Systemen, die ähnlich wie das menschliche Gehirn mentale Modelle erstellen können. Solche Systeme würden nicht nur auf historischen Daten basieren, sondern könnten hypothetische Szenarien erstellen und diese bewerten. Jedoch ist hierfür ein gutes Weltmodell notwendig, mit dem man planen kann, indem man hypothetische Abläufe modelliert. Dieses Modell muss allerdings ein abstraktes Modell der Welt sein, da man nicht jedes einzelne Luftmolekül oder Staubkorn modellieren will. Je höher die Abstraktion, desto einfacher wird das Planen, da man weniger Schritte hat, wobei jeder Schritt zu einem abstrakten Teilziel führt.

In einer Zusammenarbeit zwischen subsymbolischer KI und symbolischer KI könnte hypothetisches Denken eine entscheidende Rolle spielen. Während symbolische Systeme präzise Berechnungen und Analysen durchführen, könnten subsymbolische Systeme ihre Fähigkeit einbringen, gut in Bild- und Sprachverarbeitung zu sein, Texte gut verstehen zu können und Szenarien zu entwerfen. Diese Symbiose würde nicht nur die Effizienz steigern, sondern auch die Kreativität und Problemlösungskompetenz von KI auf ein neues Niveau heben.

Kollektives Wissen spielt ebenfalls eine entscheidende Rolle, wenn es um das menschliche Lernen geht. Es ermöglicht uns, Informationen und Erfahrungen von einer Generation zur nächsten weiterzugeben und so eine kontinuierliche Weiterentwicklung zu sichern. Das meiste Wissen, das ein Mensch in unserer heutigen Gesellschaft lernt, stammt von Erfahrungen

früherer Generationen, dem kollektiven Wissen. Kinder lernen wesentliche Dinge von den Eltern, von anderen Erwachsenen, im Kindergarten oder in der Schule. Nicht jeder Mensch muss von Neuem herausfinden, welche Pilze giftig sind und welche nicht, wie man rechnet und vieles mehr. In diesem Zusammenhang wurde der Begriff »Wissensgesellschaft« geprägt, in der individuelles und kollektives Wissen vermehrt werden.

Die menschliche Fähigkeit, Wissen zu speichern und weiterzugeben, ist untrennbar mit der Entwicklung von Sprache, Schrift und Medien verbunden. Diese Werkzeuge haben es möglich gemacht, komplexe Ideen und Technologien über Jahrhunderte hinweg zu bewahren und zu verbessern. Ein Beispiel dafür ist die Weitergabe wissenschaftlicher Erkenntnisse über Bücher, die es uns erlaubt, auf den Errungenschaften vorheriger Generationen aufzubauen. Nur wenig Wissen wird von einzelnen Menschen neu generiert. Um neues Wissen zu erlangen, gibt es die Wissenschaft und die Forschung. Hier wird unser Wissensschatz kontinuierlich vergrößert, organisiert und gespeichert.

Im Gegensatz dazu beginnen KI ihr Lernen stets bei null und sind auf riesige Mengen an Trainingsdaten angewiesen, um Wissen zu erwerben. KI greifen in der Regel nicht auf Wissen zurück, das eine andere KI gelernt hat. Jedoch werden »Foundation Models« immer beliebter, die auf großen Datensätzen aufbauen. Dadurch hat man Basismodelle erzeugt, die sehr viel Wissen in einer Domäne gelernt haben.

Domänen können allgemeiner Text, Luft- und Flüssigkeitsströmungen, Körner oder Mischungen von Mikropartikeln sein. Die Foundation Models lernen allgemeine Regeln und Gesetze der Domäne, etwa Grammatik, Semantik und Objektbeziehungen bei Text, physikalische Grundprinzipien und Teilcheninteraktionen bei Luftströmung und Körnern. Hat man die allgemeinen Prinzipien gelernt, kann man diese Modelle an spezielle Aufgaben anpassen. Das heißt, sie werden auf spezielle neue Aufgaben eingestellt, indem sie an diesen lernen. In der Anpassungsphase braucht man dann nur wenige Daten, um beispielsweise bei Text das Modell an die Automobilindustrie oder die Medizin anzupassen oder die

Oberflächenströmung eines Formel-1-Wagens zu erlernen. Das Lernen von großen Domänen-Datensätzen könnte man analog zum kollektiven Wissen sehen, das dann zur Lösung einer speziellen Aufgabe verwendet wird.

»Large Language Models« (LLMs) lernen alles menschliche Wissen, das in Textform vorhanden ist – das kollektive Wissen der Menschheit. Jedoch fehlt ihnen noch Weltverständnis, logisches Schließen, Planungsfähigkeiten und Abstraktionsfähigkeiten, um dieses Wissen einsetzen zu können. Ein weiteres Problem ist das »Halluzinieren«, bei dem LLMs sich Wissen ausdenken oder anfangen zu fantasieren, wenn Fakten zu einer bestimmten Wissensfrage nicht gelernt wurden. Das bedeutet, dass man bei LLMs nicht immer genau weiß, ob es sich um faktisches Wissen handelt oder nur um etwas, das frei erfunden wurde.

Menschen zeichnen sich durch Kreativität, Empathie und Intuition aus. Maschinen hingegen punkten mit Geschwindigkeit, Genauigkeit und der Verarbeitung großer Datenmengen. Die Zusammenarbeit von Mensch und Maschine könnte in hybriden Arbeitsmodellen Gestalt annehmen. In solchen Modellen würden KI-Systeme repetitive oder datenintensive Aufgaben übernehmen, während Menschen sich auf kreative und strategische Aufgaben konzentrieren. Dies würde nicht nur die Produktivität steigern, sondern auch die Arbeitszufriedenheit erhöhen, da Menschen sich auf Aufgaben konzentrieren können, die ihnen Sinn und Erfüllung bringen. Die Kombination dieser jeweiligen Eigenschaften könnte in Bereichen wie Medizin, Bildung und Klimaforschung bahnbrechende Fortschritte ermöglichen. Beispielsweise könnten KI-Systeme komplexe Datenanalysen durchführen, während Menschen die Ergebnisse interpretieren und ethische Entscheidungen treffen.

Trotz der Potenziale einer Mensch-KI-Zusammenarbeit bleiben Herausforderungen bestehen. Die Frage nach Transparenz, Datenschutz und Verantwortlichkeit ist zentral, um Vertrauen in KI-Systeme zu schaffen. Ebenso wichtig ist, dass der Mensch immer die Kontrolle behält und die Technologie als Werkzeug dient, nicht als Ersatz für den Menschen. Trotz der beeindruckenden Fortschritte im Bereich der KI bleiben zentrale Fragen

offen: Kann eine Maschine jemals wirklich »verstehen«? Wie lassen sich Abstraktion und abstrakte Kontextualität in maschinelle Systeme übertragen? Und wo liegen die ethischen Grenzen bei der Entwicklung immer leistungsfähigerer Technologien? Diese Fragen werden die Debatte über KI und ihre Rolle in unserer Gesellschaft weiterhin prägen.

Die Zusammenarbeit von Mensch und Maschine könnte eine neue Form der kooperativen Intelligenz hervorbringen. Diese würde nicht nur die Stärken beider Seiten vereinen, sondern auch neue Möglichkeiten der Problemlösung und Innovation erschließen. Gemeinsam könnten Mensch und KI eine bessere Welt gestalten. Die Interaktion zwischen menschlicher und künstlicher Intelligenz könnte die Schwelle zu einer neuen Ära sein. Die Unterschiede und Gemeinsamkeiten beider Systeme offenbaren Potenziale, die weit über die aktuellen Anwendungsmöglichkeiten hinausgehen.

Intelligenz ist mehr als reine Rechenleistung oder die Fähigkeit, Muster zu erkennen. Sie umfasst die Fähigkeit zu Anpassung, Planung, Abstraktion und Lernen aus wenigen Beispielen. Diese Eigenschaften sind in der menschlichen Intelligenz tief verwurzelt und könnten als Leitbild für die Weiterentwicklung von KI dienen.

Die Zukunft der KI liegt aber nicht in der Nachahmung menschlicher Intelligenz, sondern in der Schaffung von Systemen, die menschliche Stärken ergänzen und erweitern. Eine solche KI würde nicht nur Werkzeuge bereitstellen, sondern als Partner agieren, der die Fähigkeiten und Potenziale des Menschen verstärkt. Der Weg dorthin erfordert nicht nur technologische Innovation, sondern auch eine klare Vision davon, welche Rolle Intelligenz – in all ihren Facetten – in unserer Zukunft spielen soll.

4

Was kommt nach ChatGPT?

Der Hype um KI hat sich in den letzten Jahren stark auf Anwendungen wie Modelle zur Textverarbeitung und -generierung sowie zur Bildverarbeitung und -generierung konzentriert. Doch diese Technologien, so beeindruckend und nützlich sie auch sein mögen, sind nur ein kleiner Teil dessen, was KI bewirken kann. Natürlich führt kein Weg daran vorbei, die bedeutenden Leistungen von Modellen wie ChatGPT anzuerkennen. Unbestreitbar hat die Skalierung von Modellen beeindruckende Ergebnisse geliefert. Große Modelle können Texte schreiben, die oft kaum von menschlichen Autoren zu unterscheiden sind. Doch was passiert, wenn die vorhandenen Datenquellen ausgeschöpft sind?

Schon jetzt gibt es Anzeichen dafür, dass größere Modelle nicht automatisch besser werden. Ein Beispiel dafür ist die mangelnde Fähigkeit zum logischen Denken. Modelle wie ChatGPT sind hervorragend darin, bekannte Logikprobleme zu lösen, aber sie scheitern oft, wenn die Probleme leicht modifiziert oder mit unwichtigen Details erweitert werden. Zudem können sie Lösungen nicht auf andere Kontexte übertragen.

Die Fähigkeit, riesige Datenmengen zu verarbeiten und daraus Muster abzuleiten, hat oft den Eindruck erweckt, dass KI-Systeme »intelligent« seien. Doch diese Intelligenz ist in vielerlei Hinsicht begrenzt. Was Modelle wie ChatGPT tatsächlich leisten, ist eine bemerkenswerte Wissensrepräsentation. Sie können Texte speichern, abrufen und in beeindruckender Weise kombinieren. Aber sie verstehen diese Texte nicht im menschlichen Sinne. Die meisten Aufgaben, die KI heute meistert, sind in gewisser Weise Wiederholungen der beim Training gesehenen Texte. Sie beruhen auf Texten, die bereits vorhanden sind, die dann bei der Ausgabe

kombiniert werden. Doch echte Intelligenz – die Fähigkeit, völlig neue Konzepte zu schaffen, Probleme auf unerwartete Weise zu lösen und die Welt wirklich zu verstehen – besitzen sie bislang nicht.

Zentrale Fragen der aktuellen KI-Forschung sind: Wie lernen Maschinen effizient aus Daten? Welche Modelle eignen sich für unterschiedliche Problemstellungen? Und wo liegen ihre Stärken und Schwächen?

Jahrzehntelang waren LSTMs führend in Bereichen wie Spracherkennung, Textanalyse, Textgenerierung und maschineller Übersetzung. Doch mit der Einführung von Transformer-Architekturen wurden sie zunehmend ersetzt, da Transformer, wie bereits erwähnt, effizienter an großen Datenmengen lernen können. Die Fähigkeit, parallele Berechnungen durchzuführen, hat Transformer-Architekturen daher zu den dominanten Modellen in Bereichen wie Sprach- und Textverarbeitung, Textgenerierung, Bilderkennung, Bildgenerierung etc. gemacht. Aber wie jede Technologie haben auch Transformer ihre Schwächen. Ihre Komplexität und der hohe Rechenaufwand machen sie für bestimmte Aufgaben unpraktisch. Zugleich wächst die Skepsis, dass sich bedeutende Fortschritte durch reine Skalierung erzielen lassen. Ein aktuelles Diskussionsthema in der KI-Community ist die Frage, ob wir das Ende der sogenannten »Scaling Laws« erreicht haben. Diese Gesetze beschreiben, wie größere Modelle mit mehr Daten und Rechenleistung immer leistungsfähiger werden. Es gibt Anzeichen dafür, dass diese Entwicklung an eine Grenze stößt. »Deep Learning scheint gegen eine Wand zu laufen«, schrieb etwa der frühere OpenAI-Forscher Ilya Sutskever, einer der Entwickler von ChatGPT, auf der Plattform X. Stattdessen könnten wir uns in eine neue Phase der KI-Entwicklung bewegen, in der Verbesserungen der Modelle weniger von Rechenleistung und mehr von intelligenten Ansätzen abhängen.

Mit xLSTM haben wir nun ein verbessertes Konzept vorgeschlagen, das die Stärken von LSTMs mit jenen der Transformer-Modelle verbindet. Die Stärke von xLSTM liegt in der linearen Rechenkomplexität in Bezug auf die Sequenzlänge, wenn in der Anwendung Daten verarbeitet oder generiert

werden. Das bedeutet, die Anzahl der Berechnungen wächst proportional zur Länge der Eingabesequenz oder generierten Sequenz. Bei Transformern wächst die Zahl der Berechnungen dagegen quadratisch – ein gewaltiger Unterschied bei der Effizienz.

Besonders in Anwendungen, bei denen Echtzeit-Verarbeitung gefragt ist, wie etwa bei Simulationen, in der Übersetzung von Texten oder beim autonomen Fahren, könnte xLSTM einen Unterschied machen. Auch in ressourcenbegrenzten Systemen wie autonomen Produktionsanlagen, Smartphones oder bei Sensoren an Geräten bietet die Architektur erhebliche Vorteile, etwa den geringeren Energieverbrauch, da weniger berechnet werden muss, oder den kleineren Speicher, der benötigt wird.

xLSTM versus Transformer

Ein herausragendes Merkmal des xLSTM ist seine Fähigkeit, globale Zustände zu speichern und sie mit aktuellen Daten zu kombinieren. In der Spracherkennung wird dies besonders deutlich: Wenn das Modell erkennt, dass ein Sprecher einen bestimmten Dialekt hat, kann es diese Information nutzen, um die Aussprache bestimmter Wörter besser zu interpretieren. Diese Fähigkeit, globale Zustände zu berücksichtigen, macht das xLSTM in vielen Anwendungen besonders leistungsstark.

Eine weitere interessante Eigenschaft des xLSTM ist seine optimierte Nutzung von GPUs (Graphics Processing Units), die für KI-Modelle verwendeten, leistungsfähigen Grafikprozessoren. Es kann Daten so aufteilen, dass sie perfekt auf die schnellen Speicher der GPUs abgestimmt sind. Dies erlaubt eine effiziente Verarbeitung, ohne dass große Blöcke von Daten auf einmal bearbeitet werden müssen. Diese Eigenschaft macht das xLSTM besonders attraktiv für Anwendungen, bei denen die verfügbare Hardware begrenzt ist oder die Rechenleistung optimiert werden muss.

Ein zentraler Unterschied zwischen xLSTM und Transformern liegt in der Art und Weise, wie sie mit Kontextinformationen umgehen. Das

xLSTM arbeitet sequenziell und baut auf dem auf, was zuvor gespeichert wurde. Wir lesen ein Buch Wort für Wort, speichern wichtige Informationen im Gedächtnis und nutzen diese, um neue Inhalte zu verstehen. Dabei fassen wir Details oft zusammen oder lassen sie weg, wenn sie für das Verständnis der Handlung nicht entscheidend sind. Der Transformer hingegen betrachtet jedes Mal die gesamte bisher gelesene Sequenz. Er greift bei jedem neuen Wort auf die vollständige Sequenz zurück und vergleicht das neue Element mit jedem vorherigen. Dies erlaubt ihm, auch kleinste Details exakt zu rekonstruieren, ist aber rechenintensiv. In gewisser Weise gleicht der Transformer jemandem, der bei jedem neuen Wort wieder von vorne beginnt, das Buch zu lesen, um sicherzugehen, dass nichts übersehen wird.

Die besondere Stärke des xLSTM ist seine Fähigkeit zur Abstraktion. Es kann Informationen zusammenfassen und konzentriert sich auf das Wesentliche. Der Transformer ist hingegen besonders stark, wenn es darum geht, präzise zu wissen, welche Sequenzelemente (Wörter) schon da waren. Bei Aufgaben, bei denen es auf chronologische Abfolgen oder zeitliche Zusammenhänge ankommt, hat das xLSTM oft die Nase vorn. Beispiele hierfür sind Zeitreihenanalysen, wie sie in Aktienkursvorhersagen, Prozesskontrolle oder Marketing genutzt werden. In vielen dieser Aufgaben hat der aktuelle Zustand eines Systems mehr Einfluss auf die nächste Entscheidung als Zustände aus der fernen Vergangenheit. Das xLSTM ist hier besonders gut geeignet, weil es sehr leicht die jüngsten Daten stärker gewichten und ältere abschwächen kann.

Eine zentrale Komponente des xLSTM ist das sogenannte »exponentielle Gating«. Um zu verstehen, warum diese Innovation so bedeutend ist, müssen wir zunächst einen Blick auf die Funktionsweise des ursprünglichen LSTM werfen. Dort wurden Informationen durch Gates gesteuert, die mithilfe der sogenannten »Sigmoid-Funktion« arbeiten. Diese Funktion liefert Werte zwischen 0 und 1: Ein Wert von 0 bedeutet, dass die Information blockiert wird, ein Wert von 1, dass sie vollständig weitergegeben wird. Der Nachteil der Sigmoid-Funktion liegt in ihrer

Begrenzung. Wenn eine Information einmal mit einem bestimmten Wert gewichtet wurde, kann diese Gewichtung später nicht mehr verändert werden. Das Problem dabei: Wenn später ein Wert kommt, der viel höher gewichtet werden soll als frühere, geht das nicht, da das Gewicht mit 1 begrenzt ist. Das originale LSTM kann Entscheidungen, das heißt Gewichtungen, nicht mehr revidieren. Das exponentielle Gating löst dieses Problem, indem es die Gewichtung nach oben hin öffnet. Die Exponentialfunktion ist nicht beschränkt, was bedeutet, dass Informationen jederzeit neu bewertet und sogar mit einem sehr hohen Faktor gewichtet werden können, falls sie später als besonders wichtig erkannt werden. So kann das xLSTM Entscheidungen, die zuvor getroffen wurden, revidieren und an neue Gegebenheiten anpassen.

Ein Beispiel macht dies deutlich: Stellen Sie sich vor, Sie suchen nach dem Preis der besten Badezimmerfliesen. Zu Beginn finden Sie ein Modell, das 10 Euro kostet, und speichern sich den Preis. Später entdecken Sie jedoch Fliesen, die nur 5 Euro kosten, und erkennen, dass diese eine bessere Wahl sind. Sie speichern sich nun 5 Euro. Noch später stoßen Sie auf ein Modell, das 100 Euro kostet, aber von überragender Qualität ist. Nun sollen Sie 100 Euro speichern, was das originale LSTM nicht kann. Jedoch kann xLSTM mit exponentiellem Gating die ursprüngliche Entscheidung (10 Euro) vergessen und die neue (100 Euro) stärker gewichten.

Ein weiterer Meilenstein des xLSTM ist die Einführung eines Matrix-Speichers. Im klassischen LSTM wurde die Information in einem skalaren Wert gespeichert, also einer einzigen Zahl. Diese Zahl repräsentierte beispielsweise, wie oft in einem Text Wörter zu einem bestimmten Thema wie »Sport« oder »Katzen« vorkamen. Diese Art der Speicherung war zwar effizient, stieß jedoch bei komplexeren Zusammenhängen schnell an ihre Grenzen. Vorteil dieser Speicherung war, dass dies zu guten Abstraktionen führte, die LSTM stark machten, aber für die heutige Textgenerierung nicht ausreichend waren. Der Matrix-Speicher des xLSTM bringt hier eine deutliche Verbesserung. Das xLSTM kann nicht nur eine Zahl speichern, sondern ganze Vektoren – also mehrere Werte gleichzeitig, die gemeinsam

ein komplexes Muster beschreiben. Diese Vektoren ermöglichen es dem Modell, präzisere und detailliertere Informationen zu behalten. In der Praxis bedeutet das, dass das xLSTM nicht nur erkennt, dass in einem Text von einem »Fußballspieler« die Rede ist, sondern auch, dass es sich etwa um »Manuel Neuer« handelt, der eine rote Karte bekommen hat. Dieses Wissen hilft, präzisere Texte zu generieren. Wenn im Text etwa ein »Fußballer« erwähnt wird, der eine rote Karte bekommen hat, dann weiß das xLSTM, dass es eben um Manuel Neuer geht.

Der Matrix-Speicher basiert auf Konzepten, die bereits in den 1970er-Jahren entwickelt wurden und später in den 1980ern mit den sogenannten »Hopfield-Netzwerken« popularisiert wurden. Diese Netzwerke waren darauf ausgelegt, Muster zu speichern und wiederzuerkennen. Das xLSTM erweitert diese Idee, indem es die Vorteile des Hopfield-Matrix-Speichers mit den Mechanismen des LSTM kombiniert. Die LSTM-Gates erweitern den Speicher und bestimmen nun, was in den Speicher geschrieben wird und wann Informationen aus dem Speicher abgefragt werden. Dadurch entsteht ein leistungsfähiger Speicher, der sowohl präzise als auch flexibel ist.

Das dritte große Feature des xLSTM ist seine bessere Parallelisierbarkeit. Das bedeutet, dass das Modell Daten schneller verarbeiten kann, weil es Informationen nicht sequenziell, sondern gleichzeitig verarbeitet. Dieser Ansatz ist besonders bei Aufgaben wie der Verarbeitung natürlicher Sprache oder großen Datensätzen von Vorteil, bei denen Geschwindigkeit entscheidend ist. Die Parallelisierbarkeit bezieht sich hier nur auf das Lernen, wenn die zu lernende Sequenz schon vollständig vorliegt. Hier kann die gesamte Sequenz gleichzeitig verarbeitet werden, das heißt, die Sequenz wird parallel verarbeitet. In der Anwendung, der Inferenz, ist die parallele Verarbeitung üblicherweise nicht möglich. Erzeugt man Texte, dann gibt es keine Parallelität, da ja die Ausgabesequenz Schritt für Schritt, also Wort für Wort, erzeugt wird. Man muss ein Wort nach dem anderen sequenziell erzeugen, was den Transformer in der Anwendung so langsam macht, da er ja bei jedem neuen

Wort auf die ursprüngliche Frage und jedes schon erzeugte Wort zurückschaut.

Im klassischen LSTM ist das Lernen im Training nicht parallelisierbar, da der neue Speicherzustand von dem alten abhängt. Das bedeutet, dass Informationen sequenziell verarbeitet werden müssen – also erst muss der alte Speicherzustand berechnet werden, bevor der neue berechnet werden kann. Diese Abhängigkeit schränkt die Geschwindigkeit ein und macht paralleles Arbeiten unmöglich. Das xLSTM löst dieses Problem, indem es eine Abhängigkeit zwischen den Speicherzuständen vermeidet. Neue Informationen werden unabhängig von den bereits gespeicherten verarbeitet und dann dem Speicher hinzugefügt. Dadurch kann das Modell mehrere Speicheroperationen gleichzeitig durchführen, was die Verarbeitungsgeschwindigkeit erheblich steigert. Dies ist besonders für Anwendungen wichtig, bei denen große Datenmengen gelernt werden müssen, wie etwa in der Verarbeitung von natürlichsprachlichem Text.

Die Neuerungen des xLSTM – exponentielles Gating, Matrix-Speicher und Parallelisierbarkeit – eröffnen neue Möglichkeiten für den Einsatz von xLSTM. Sie machen das xLSTM nicht nur flexibler und leistungsfähiger als seinen Vorgänger, das LSTM, sondern ermöglichen es auch, komplexere Probleme zu lösen.

Ein Beispiel ist die Verarbeitung von Texten. Mit dem xLSTM können Zusammenhänge zwischen Wörtern besser erkannt und gespeichert werden als mit dem originalen LSTM, was Anwendungen wie maschinelle Übersetzungen, Textzusammenfassungen oder Textgenerierung verbessert.

Auch in der Bildverarbeitung war das neue xLSTM sehr erfolgreich durch ViL (Vision-xLSTM), das bessere Ergebnisse liefert als Vergleichsmethoden, die meist auf der Transformer-Architektur basieren. So brillierte das ViL etwa bei der semantischen Segmentierung (»Vision-LSTM: xLSTM as Generic Vision Backbone«). Ein Beispiel: Bei der semantischen Segmentierung bei der Bildverarbeitung für selbstfahrende Autos müssen alle Pixel eines Fußgängers, eines Fahrradfahrers, eines Pkw, eines Lkw, einer Straße, einer Ampel, eines Verkehrsschildes oder einer

Straßenmarkierung zusammengefasst werden, damit das Objekt mit Position und Größe weiterverarbeitet werden kann. Das Vision-xLSTM ist nicht nur genauer, sondern auch noch schneller als andere Methoden, die beispielsweise auf der Transformer-Architektur beruhen.

Das xLSTM zeigt seine Stärke auch bei anderen Anwendungen wie der Verarbeitung von Zeitreihen. Zeitreihen, wie sie in der Vorhersage von Überschwemmungen (Hydrologie), Aktienkursvorhersagen, Prozesskontrolle (z. B. Kernfusion) oder im Marketing vorkommen, erfordern ein Modell, das die zeitliche Abfolge der Daten versteht. Das xLSTM ist hierfür ideal geeignet, da es die aktuellen Zustände stärker gewichtet und ältere Informationen automatisch abschwächt. Ein Beispiel aus der Hydrologie: Der aktuelle Regenfall hat oft einen größeren Einfluss auf die Vorhersage des Wasserstands in einem Fluss als der Regenfall vor mehreren Stunden. Das ist wichtig, um Überschwemmungen genau vorherzusagen. Ähnlich verhält es sich in Produktionsanlagen, bei denen chemische Prozesse überwacht werden. Hier müssen Parameter wie Temperatur oder Druck kontinuierlich erfasst und in Beziehung zueinander gesetzt werden, um optimale Ergebnisse zu erzielen. Das xLSTM kann nicht nur den aktuellen Zustand speichern, sondern auch globale Parameter berücksichtigen. Die Fähigkeit, globale Zustände zu speichern und mit den aktuellen Daten zu kombinieren, macht das xLSTM in vielen Anwendungen zur besten Methode.

Ein weiteres Beispiel ist die Analyse von Aktienkursen: Das xLSTM analysiert die Kursdaten Zeitschritt für Zeitschritt, speichert relevante Informationen im Gedächtnis und nutzt diese, um zukünftige Kurse vorherzusagen. Während das xLSTM für eine neue Vorhersage nur das Gedächtnis mit den aktuellen Daten aktualisieren muss, muss der Transformer bei einer neuen Vorhersage wieder über alle bisherigen Kursdaten gehen, was viel rechenintensiver ist.

In den Lebenswissenschaften hat das xLSTM ebenfalls vielversprechende Anwendungen gefunden: In der DNA-Sequenzanalyse zum Beispiel müssen lange Sequenzen analysiert und Muster erkannt

werden, die entscheidend für das Verstehen von Krankheiten sind. Das xLSTM kann nicht nur die aktuelle Sequenz verarbeiten, sondern auch Abhängigkeiten von verschiedenen Stellen in der DNA erkennen. Dazu muss man wissen, dass die DNA als Sequenz dargestellt ist, aber in der menschlichen Zelle eine 3D-Struktur hat. Deshalb können Stellen, die in der Sequenz weit voneinander getrennt sind, in 3D miteinander interagieren.

Auch in der Robotik hat das xLSTM klare Vorteile in Sachen Speichereffizienz und Schnelligkeit. Ein Roboter muss seinen aktuellen Zustand – beispielsweise die Position eines Arms – genau kennen, um die nächste Bewegung korrekt auszuführen. Ohne diese Zustandsinformationen müsste er den gesamten Prozess nochmals neu durchrechnen, was ineffizient wäre. Das xLSTM kann diese Zustände speichern und aktualisieren, was es ideal für Steuerungsaufgaben macht. Dank seiner Fähigkeit, globale Zustände zu speichern und zu berücksichtigen, bietet es auch in komplexen Anwendungen eine robuste Lösung. Ein noch viel größerer Vorteil ist die Schnelligkeit von xLSTM bei der Anwendung, was für die Robotik essenziell ist. Die Bewegungen sollen fließend sein und nicht ruckelnd, weil auf Berechnungen gewartet werden muss.

Insbesondere bei der Herstellung individualisierter Produkte können xLSTM-gesteuerte Roboter eine wichtige Rolle spielen. Sie lernen durch wenige Beispiele im Kontext. Das heißt, statt einer Frage werden bei der Aufgabenstellung Beispiele gegeben und gesagt, dass es nun die folgenden Aufgaben genauso lösen soll. Beispielsweise kann man ein Lackierungs-Foundation-Modell trainieren, das eine riesige Anzahl von Lackierungsproblemen im Training gesehen hat. Nun zeigt man dem Modell im Kontext wenige Beispiele, wie ein neuer Kotflügel lackiert werden soll. Es soll dann weitere Kotflügel der gleichen Serie allein lackieren können. Solche Modelle können so flexibel auf neue Anforderungen reagieren, da Beispiele nicht gelernt, sondern als Kontext gegeben werden. Das würde die Produktion von Kleinserien wirtschaftlicher machen, denn das Programmieren von Industrierobotern für Kleinserien lohnt sich nicht. Wenn man aber nur wenige Beispiele, also

Demonstrationen, braucht, damit der Roboter lernt, was er tun soll, kann man Roboter auch für Kleinserien einsetzen.

Die Entscheidung, ob ein xLSTM oder ein Transformer eingesetzt werden sollte, hängt von der spezifischen Aufgabe ab. Zeitreihen, Robotik und Spracherkennung sind typische Anwendungsfelder des xLSTM. Der Transformer hingegen brilliert in Aufgaben, bei denen Sequenzelemente wie Wörter exakt gespeichert werden müssen. Das ist typischerweise bei Fragen zu einem Text der Fall, wo erst ein Text vorgegeben wird und dann eine Frage dazu gestellt wird. Beispielsweise wird erst das Buch *Der Hobbit* gelesen und dann gefragt: Wo hat Bilbo den Ring versteckt? Dies geht allerdings auch mit xLSTM. Man müsste nur zuerst die Frage stellen und ihm dann das Buch geben. Dann wird aus der Gedächtnisaufgabe eine Suchaufgabe: Das System muss nur nach der richtigen Stelle im Buch suchen, was xLSTM gut kann. Mit dem neuen Memory von xLSTM zeigte sich, dass es mit Transformern bei der Generierung von Texten mithalten kann. Denn auch dort ist es wichtig, sich die wenigen relevanten Wörter zu merken.

Eine zentrale Frage bei neuronalen Netzen ist, wie schnell sie lernen und wie viele Daten sie dafür benötigen. Durch die Analyse mit Scaling Laws stellte sich heraus, dass xLSTM dem Transformer immer überlegen ist. Bei einer fest vorgegebenen Rechenzeit kann man immer ein xLSTM-Modell trainieren, das leistungsfähiger ist als alle Transformer-Modelle, die mit der gleichen Rechenzeit trainiert wurden.

Umgekehrt: Bei einer fest vorgegebenen Leistung, zum Beispiel bei der Textanalyse, kann man immer ein xLSTM-Modell trainieren, das weniger Rechenzeit braucht als alle Transformer-Modelle, die die gleiche Leistung haben wie das xLSTM-Modell.

Das Gleiche gilt für Daten: Bei einer fest vorgegebenen Trainingsdatenmenge kann man immer ein xLSTM-Modell trainieren, das leistungsfähiger ist als alle Transformer-Modelle, die mit dem gleichen Trainingsdatensatz trainiert wurden. Bei einer fest vorgegebenen Leistung, etwa bei der Textanalyse, kann man immer ein xLSTM-Modell trainieren,

das weniger Trainingsdaten braucht als alle Transformer-Modelle, die die gleiche Leistung haben wie das xLSTM-Modell.

Natürlich hängt die Leistung beider Modelle von der Anpassung der jeweiligen Lernparameter ab, weshalb in allen Vergleichen die Lernparameter der Modelle zuerst optimiert werden.

In Bezug auf die Modellgröße und Datenanforderungen zeigt sich, dass das xLSTM effizienter ist. Im Gegensatz zu Transformern generalisiert das xLSTM auch, wenn wenig Daten zum Lernen vorhanden sind.

Das xLSTM bietet zudem einen großen Vorteil in der Effizienz. Durch seine Fähigkeit, Daten rekurrent und sequenziell zu verarbeiten, kann es Hardware-Ressourcen wie GPUs optimal nutzen. Es erlaubt, Daten in kleinen Häppchen (Chunks) zu verarbeiten, die perfekt auf die Kapazität der Hardware abgestimmt sind. Transformer hingegen müssen mit großen Blöcken arbeiten, was die Anpassung an spezifische Hardware erschwert und mehr Energie erfordert. Die neuen xLSTM-Hardware-Implementationen sind im Training schneller als die Transformer-Implementationen, die auf der berühmten FlashAttention-Methode aufbauen. Eine Kuriosität des Schicksals: Transformer haben LSTM bei der Textverarbeitung und -generierung verdrängt, da sie schneller im Training waren. Nun ist xLSTM im Training schneller als Transformer. Und xLSTM ist natürlich schneller in der Anwendung, da es eine lineare Rechenkomplexität hat, im Gegensatz zur quadratischen des Transformers.

Ein weiterer Vorteil des xLSTM ist seine Fähigkeit, diese Verarbeitung modular zu gestalten. Es kann effizient zwischen sequenzieller Verarbeitung und paralleler Verarbeitung wechseln, je nachdem, ob es sich in der Trainingsphase befindet oder in der Inferenz, d. h. der Anwendung. Dies ermöglicht eine flexible Nutzung in verschiedenen Szenarien und steigert sowohl die Trainingsgeschwindigkeit als auch die Effizienz bei der Inferenz.

Geheimnis des Erfolgs von Deep Learning: Lösungspfade zu guten Modellen

In der KI-Forschung stoßen wir immer wieder auf das Problem, dass unsere Modelle oft in gewisser Weise »tunnelblickartig« agieren. Sie fokussieren sich auf das, was im ersten Moment am vielversprechendsten erscheint, und ignorieren dadurch möglicherweise bessere, aber zunächst weniger offensichtliche Lösungen. Dieses Phänomen begegnet uns in vielen Bereichen, sei es in der Bildverarbeitung oder bei der Sprachverarbeitung.

Unsere heutigen KI-Modelle basieren fast ausschließlich auf gradientengestütztem Lernen wie dem Backpropagation. Sie optimieren ihre Leistung, indem sie den Fehler schrittweise in Richtung des Gradienten minimieren. Der Gradient gibt die Richtung der Parameter an, in der sich der Fehler am stärksten reduziert. Doch das führt dazu, dass sie immer nur den direktesten, steilsten Weg der Fehlerabnahme nehmen. Dabei übersehen sie alternative Pfade, die vielleicht zunächst den Fehler nicht so stark verkleinern, aber langfristig zu besseren Ergebnissen führen könnten. Die Ursache des Übersehens von alternativen Pfaden liegt im Backpropagation-Algorithmus, der versucht, den Fehler möglichst schnell klein zu machen. Dadurch konzentriert sich Backpropagation auf die erfolgversprechendsten Muster und übersieht oft alternative Muster, die jedoch in bestimmten Situationen sehr nützlich sein können.

Ein Beispiel aus der Bildverarbeitung: Systeme lernen zunächst einfache Merkmale wie Texturen, weil diese leichter lernbar sind und schnell zu einer Reduktion des Fehlers führen. Die Umrisse oder Formen eines Objekts – wie der Umriss eines Tigers – werden oft nicht gelernt, weil dieser Weg den Fehler nicht schnell minimiert. Als Menschen erkennen wir einen Tiger auch dann, wenn er nur als schwarze Silhouette dargestellt ist, aber ein KI-System versagt hier, weil es auf Texturen wie Fell oder Augen trainiert wurde. Die Textur eines Tigers kann Schritt für Schritt gelernt werden, da der Fehler schrittweise verkleinert werden kann: Erst werden die typischen Farben gelernt, dann Farbübergänge (Farbkanten), dann die Streifen usw. Jedoch ist die Silhouette eines Tigers schwer Schritt für Schritt lernbar, da Teile des Umrisses nichts nützen und der Umriss immer wieder anders aussieht, je nachdem, wie der Tiger sich bewegt.

Ich vergleiche das gern mit dem Lesen eines Buches. Wenn wir einen Krimi lesen, konzentrieren wir uns instinktiv auf Elemente wie ein Messer oder eine Pistole, weil wir gelernt haben, dass solche Details später wichtig sein könnten. Wir übersehen dabei scheinbar unwichtige Hinweise, die im größeren Zusammenhang entscheidend sind. KI-Systeme verhalten sich ähnlich: Sie priorisieren die Informationen, die sofortige Ergebnisse liefern, und übersehen dabei die subtileren, aber potenziell wichtigeren Aspekte.

In den Diskussionen über KI-Modelle und ihre Fortschritte wird oft der Fokus auf die Qualität der gefundenen Lösungen gelegt. Doch mindestens genauso wichtig ist die Frage, wie diese Lösungen überhaupt gefunden werden. Es geht nicht nur darum, ob es Lösungen gibt, sondern ob es auch zugängliche Pfade dorthin gibt – Pfade, die unsere heutigen Lernmethoden entlanggehen können. Im obigen Beispiel mit dem Erkennen eines Tigers führt der Pfad mit der Textur zu einer Lösung, da die Textur Schritt für Schritt verfeinert werden kann. Der Fehler nimmt dabei Schritt für Schritt ab. Bei der Silhouette hingegen ist eine schrittweise Fehlerverbesserung viel schwieriger, da Teile der Silhouette allein nicht ausreichend sind und es je nach der Pose des Tigers viele Silhouetten gibt. Im Gegensatz dazu hat der Mensch von Kindheit an gelernt, wie Silhouetten ineinander übergehen, wenn er die Welt beobachtet.

Ein anderes Beispiel für die Wichtigkeit der Existenz eines Lösungspfades ist das schon erwähnte »Forget Gate« bei LSTM. Dieser Mechanismus soll es einem neuronalen Netz ermöglichen, irrelevante Informationen zu vergessen, während wichtige behalten werden. Doch in der Praxis zeigt sich, dass das Lernen des Vergessens ein extrem schwieriger Prozess ist. Denn um zu vergessen, muss ein Modell zuerst lernen, sich etwas zu merken. Sobald die Netze anfangen, Teile dieses Wissens zu vergessen, verschlechtert sich ihre Fähigkeit, sich etwas zu merken, denn sie vergessen ein klein wenig in jedem Schritt.

Ein LSTM kann sich Informationen merken oder vergessen, indem es das Forget Gate steuert. Wenn das Forget Gate den Wert 1 hat, wird eine Information behalten. Wenn es den Wert 0 hat, wird die Information

vergessen. Nun gibt es Situationen, in denen das LSTM etwas vergessen soll. Beispielsweise kann es ein starkes Muster an einer bestimmten Stelle in der Sequenz geben, daher ist es nun irrelevant, ob ein bestimmtes schwächeres Muster bereits da war. Das Forget Gate sollte den Speicher für das schwächere Muster an dieser Stelle löschen. Dann wird der Lernalgorithmus das Forget Gate für dieses Beispiel an dieser Stelle in Richtung null drücken, also verkleinern. Doch das Problem ist: Das Netzwerk weiß nicht genau, was genau das starke Muster ist, bei dem es das schwache Muster vergessen soll, da ein einzelnes Beispiel nicht ausreicht, um das starke Muster zu lernen. Das Forget Gate wird nicht nur bei dem starken Muster verkleinert, sondern auch für Muster, die ähnlich zum starken Muster sind. Dadurch wird das LSTM auch dort vergessen, wo es sich eigentlich genau erinnern sollte – die Merkfähigkeit verschlechtert sich und der Fehler wird größer.

Um den Fehler zu korrigieren, wird der Lernschritt zum Vergessen des schwächeren Musters rückgängig gemacht. Ein Beispiel hierzu: Ich möchte zählen, wie viele Schritte ich in verschiedenen Waldstücken mache. Immer wenn ich einen Wald betrete, fange ich zu zählen an, wenn ich den Wald verlasse, schreibe ich die Anzahl der Schritte auf einen Zettel – und kann die aktuelle Anzahl von Schritten vergessen.

Für ein LSTM ist diese Aufgabe sehr schwierig, da es lernen müsste, die Anzahl der Schritte zu vergessen, wenn ich den Wald verlasse. Also müsste es das Verlassen des Waldes erkennen und dann dort das Forget Gate auf null setzen. Beispielsweise lernt das LSTM, eine Wiese als Hinweis für das Verlassen des Waldes zu erkennen. Jedoch gibt es im Wald auch Lichtungen, Gebiete mit Graswuchs oder Blumen, die Ähnlichkeiten zu einer Wiese haben. Nun würde das Forget Gate auch bei diesen Waldgebieten mit Gras oder Blumen viel kleiner werden und die Schritte könnten nicht mehr gezählt werden. Deshalb wird das Vergessen nicht gelernt. Natürlich könnte das LSTM eine Wiese erkennen, aber nicht mit einem oder wenigen Beispielen, denn die ersten Beispiele reichen nicht, um eine Wiese von einer Lichtung zu unterscheiden, und würden so das

Gelernte, also das Zählen, verschlechtern und so größere Fehler erzeugen. Es gibt keinen Pfad zur Lösung mit Vergessen, bei dem der Fehler immer kleiner wird. Das war schon immer ein Problem des Forget Gates: Es gibt normalerweise keinen Pfad mit Fehlerabnahme zum Vergessen, da das einhergehende schlechtere Merken den Fehler groß macht.

Vergessen ist keine isolierte Fähigkeit, die einfach hinzugelernt werden kann. Es ist Teil eines komplexen Lernprozesses, bei dem jede Änderung an einer Stelle Auswirkungen auf viele andere Bereiche hat. Unsere heutigen Modelle sind nicht darauf ausgelegt, solche langfristigen und indirekten Abhängigkeiten effektiv zu bewältigen.

Dieses Verhalten ist symptomatisch für ein grundlegendes Problem in der KI: Systeme optimieren ihre Leistung so, dass kurzfristige Verbesserungen priorisiert werden. Langfristige Strategien, die zunächst zu größeren Fehlern führen, werden nicht verfolgt, weil sie kurzfristig mit einem scheinbaren Nachteil verbunden sind. Dieses Muster zeigt sich nicht nur beim Vergessen, sondern auch in anderen Bereichen des Lernens, etwa bei der Wahl von Merkmalen in der Bildverarbeitung, wie oben schon erwähnt am Beispiel des Tigers.

Ich glaube, dass das Konzept der Lösungspfade das Geheimnis des Erfolgs von Deep Learning ist. Lösungspfade sind nicht nur beim Vergessen relevant. Sie spielen eine zentrale Rolle beim Lernen von Deep-Learning-Modellen und insbesondere von großen Netzwerken. Große Netzwerke eröffnen mehr Pfade zu möglichen Lösungen. Sie ermöglichen es, neue Muster und Strukturen zu lernen, ohne dabei bereits gefundene Muster und Strukturen zu zerstören. Das ist ein entscheidender Vorteil gegenüber kleineren Netzwerken, die oft in ihren Möglichkeiten eingeschränkt sind, da sie weniger Pfade zu Lösungen haben.

Für mich ist klar, dass unseren heutigen KI-Methoden etwas Entscheidendes fehlt. Wir Menschen lernen nicht immer direkt und linear – wir sind in der Lage, Umwege zu gehen, scheinbar unwichtige Details zu erfassen und später zu erkennen, dass diese doch entscheidend sind. Unsere Gehirne funktionieren nicht nach den exakten mathematischen Prinzipien,

auf denen heutige KI-Systeme basieren. Stattdessen arbeiten sie mit Annäherungen und suboptimalen Lösungen, die sich erst im Nachhinein als sinnvoll herausstellen. Mit der Einführung großer Modelle haben wir einen wichtigen Schritt in die richtige Richtung gemacht, da es neue Pfade zu bisher unentdeckten Lösungen gibt. Aber der Mensch macht es noch besser bei der Auffindung guter Lösungen.

Ich denke, dass KI-Systeme zu sehr darauf ausgerichtet sind, ständig den Fehler zu minimieren. In vielen Fällen müssten sie jedoch zuerst den Fehler »bergauf« gehen, um später einen besseren Weg hinab in bisher ungesehene Fehlerminima zu finden. Diese Fähigkeit fehlt ihnen derzeit, und genau hier sehe ich eine der größten Herausforderungen für die Zukunft der KI. Ich glaube, dass wir mit den heutigen Methoden viele Lösungen übersehen, da wir zu gierig, aber nicht umsichtig lernen. Wir wollen sofort den Fehler möglichst klein haben, ohne zuerst wesentliche Konzepte gelernt zu haben. Wir versuchen, mit unseren neuronalen Netzen bei der Bildverarbeitung sofort Objekte zu erkennen, ohne zu lernen, was Objekte zu Objekten macht.

Menschen lernen auf andere Weise, Objekte zu erkennen. Beispielsweise wird »Visual Binding« sehr früh gelernt. Beim »Visual Binding«-Problem geht es um die Frage, warum wir einen roten Kreis und ein blaues Quadrat nicht mit einem blauen Kreis und einem roten Quadrat verwechseln. Oder welche Bereiche in einem Bild zu demselben Objekt gehören. Es können Bereiche mit gleicher Textur sein oder Bereiche, die sich synchron bewegen wie bei einem Auto. Merkmale von Objekten wie Farbe und Form werden in unserem Gehirn erst getrennt und dann wieder kombiniert. Die Trennung ermöglicht Abstraktionen und effiziente Verarbeitung, wohingegen es wichtig ist, zu wissen, welches Merkmal zu welchem Objekt gehört.

Es zeigt sich, dass das menschliche Lernen effektiver und weniger fehleranfällig ist. Das kann man an den »Adversarial Attacks« in der Bildverarbeitung festmachen. Bei Adversarial Attacks wird ein Bild so manipuliert, dass ein Mensch diese Veränderung nicht bemerkt, während das neuronale Netz jedoch die Objekte auf dem Bild falsch zuordnet,

obwohl es zuerst alles richtig gemacht hat. Der Mensch hat hier eine robustere und weniger fehleranfällige Methode der Objekterkennung gelernt. Hier fehlt der KI noch etwas, insbesondere beim Auffinden guter Modelle.

Blinde Flecken

Die Frage ist, wie wir diesen blinden Fleck bei der Lösungssuche der KI überwinden können. Es braucht neue Ansätze. Bis dahin bleibt unsere Herausforderung, die Modelle so zu steuern, dass sie über ihren Tunnelblick hinauswachsen und den größeren Zusammenhang erkennen können – etwas, das uns Menschen selbstverständlich ist.

Eine wichtige Richtung, die KI-Forschung und -Anwendung nehmen könnten, liegt im Bereich der Inferenz – also der Anwendung eines gelernten Modells. Viele der heutigen KI-Systeme enthalten bereits enorme Mengen an Informationen. Die Herausforderung besteht darin, diese Informationen abzurufen – und zwar solche, die für die Nutzer nützlich sind.

Man kann sich das wie ein Interview vorstellen: Wenn man die falschen Fragen stellt, bekommt man keine brauchbaren Antworten. Doch mit der richtigen Herangehensweise können auch bestehende Systeme deutlich besser genutzt werden. Die richtigen Fragen zu stellen, das entspricht dem richtigen Prompt für LLMs, welches oft auch eine Frage sein kann. Es hat sich sogar das Feld des »Prompt-Engineering« gebildet.

Prompt-Engineering ist die Aufgabe der Erstellung einer Anweisung, um die bestmögliche Ausgabe eines generativen KI-Modells zu bekommen. Der Prompt ist meist ein Text, der eine Frage stellt oder eine Aufgabe beschreibt, die eine KI lösen soll. Für Textgenerierung kann ein Prompt ein Text sein, der eine Frage, einen Befehl oder einen längeren Text mit Kontext und Anweisungen enthält. Oft wird ein Stil, eine Wortwahl oder ein Bezug vorgegeben. Bei Bildgenerierungs-KI wird im Prompt das gewünschte Bild

beschrieben. Ein berühmtes Beispiel von der OpenAI-KI DALL-E ist der Astronaut, der auf einem Pferd reitet. Prompts für Bildgenerierung enthalten ein gewünschtes Thema, einen Stil, Objekte, die vorhanden sein sollen, ein Layout oder eine Beleuchtung bis hin zu Kameraeinstellungen.

Im Gegensatz dazu zielt das »Chain of Thought (CoT) Prompting« bei Textmodellen mit Mathematik- oder Logikfragen darauf ab, die Leistung zu verbessern. Ein Benutzer fügt eine Anweisung wie »Beschreiben Sie Ihre Argumentation in Schritten« oder »Erklären Sie Ihre Antwort Schritt für Schritt« an das Ende eines Prompts. Diese Prompts bringen das LLM dazu, nicht nur ein Ergebnis zu generieren, sondern auch die Zwischenschritte zu beschreiben, die zu diesem Ergebnis geführt haben.

Neuere KI »denken« mehr nach, das heißt, sie investieren mehr Rechenzeit in die Inferenz. So hat OpenAI die Modelle o1 (strawberry) und o3 herausgebracht, DeepSeek das Modell R1. All diese Modelle investieren viel mehr Zeit in die Inferenz und analysieren verschiedene CoTs. Während der Inferenz werden Lösungen modifiziert, überarbeitet oder an bestimmten Stellen neue Lösungspfade beschritten. DeepSeek setzte, um ihr R1 für mathematische Aufgaben besser zu machen, Monte-Carlo-Tree-Search ein. Dabei handelt es sich um einen Algorithmus zur Entscheidungsfindung in Spielen und komplexen Problemlösungen, der zufällige Simulationen mit einer baumartigen Suche kombiniert, um die besten möglichen nächsten Schritte zu finden.

Monte-Carlo-Tree-Search war sehr erfolgreich bei AlphaGo/AlphaZero, um den Go-Weltmeister zu schlagen. Die Aufgabenstellungen für Go und Mathematik ähneln sich. Ich kann eine Go-Partie bis zum Ende durchspielen und sehen, ob ich gewinne. Andererseits kann ich einen mathematischen Lösungspfad verfolgen und sehen, ob ich am Ende die richtige Lösung bekomme. Da viel mehr Zeit in die Inferenz gesteckt wird, sehe ich hier auch eine Chance für xLSTM, das gerade in der Inferenz sehr schnell ist.

Heutige KI-Lernmethoden sind darauf ausgelegt, Lösungen durch mathematische Optimierung zu finden. Wir Menschen lernen nicht gierig,

um einen Fehler zu minimieren, und wir versuchen auch nicht immer, das Minimum oder Maximum zu finden. Wenn wir zum Mittagessen gehen, dann optimieren wir in der Regel nicht den Weg oder die Anzahl der Schritte, die wir machen. Für Menschen reicht es oft, dass wir eine Aufgabe hinreichend gut lösen, aber nicht, dass wir sie optimal lösen. KI-Systeme sind darauf ausgelegt, immer die optimale Lösung zu finden, und das von Anfang an. Deshalb nehmen KI-Systeme – wie schon gesagt – manchmal den falschen Pfad, der zu Anfang den Fehler stark minimiert, aber später zu weniger guten Lösungen führt. Hier sind wir Menschen besser.

Ich glaube nicht, dass das Gehirn auf dieselbe Weise lernt wie KI. Vielleicht liegt der Schlüssel in einem Prinzip, das wir bisher übersehen haben – ein Prinzip, das uns im Nachhinein offensichtlich erscheinen wird, aber heute noch völlig unklar ist.

Unsere Modelle haben beeindruckende Fortschritte gemacht, doch sie stoßen an Grenzen. Um die nächste Stufe der KI-Forschung zu erreichen, müssen wir nicht nur neue Lösungen finden, sondern auch neue Pfade zu diesen Lösungen schaffen. Das erfordert ein Umdenken in der Art und Weise, wie wir KI entwickeln – weg vom reinen Optimieren hin zu einer flexibleren, menschenähnlicheren Denkweise. Die Zukunft der KI hängt davon ab, ob wir diesen Paradigmenwechsel vollziehen können.

5

Die Industrialisierung der KI

Die Entwicklung der Gesellschaft ist untrennbar mit der Erfindung bahnbrechender, disruptiver Technologien verbunden. Jede neue, transformative Technologie durchläuft dabei drei zentrale Phasen: Grundlagenforschung, Skalierung und Industrialisierung. Die drei Phasen verdeutlichen, wie sich eine Technologie von den ersten wissenschaftlichen Entdeckungen bis zur vollen Integration in die Wirtschaft und die Gesellschaft entwickelt. Dieses Drei-Phasen-Konzept kann uns helfen, die Entwicklung der modernen KI, also des Deep Learnings, besser zu verstehen.

Historische Beispiele, anhand derer sich die drei genannten Phasen gut nachvollziehen lassen, sind etwa die Dampfmaschine und der Haber-Bosch-Prozess. Die Dampfmaschine, die die industrielle Revolution einleitete, durchlief eine Phase der experimentellen Grundlagenforschung, bevor sie durch Skalierung leistungsfähiger wurde und schließlich die industrielle Produktion und den Transport grundlegend veränderte. James Watt war ein Forscher, der die Industrialisierung der Dampfmaschine durch Nutzen der Dampfexpansion mit Kolben und dem Fliehkraftregler vorantrieb.

Der Haber-Bosch-Prozess, der die synthetische Herstellung von Ammoniak als Düngemittel ermöglichte und die Grundlage moderner Landwirtschaft schuf, folgte einem ähnlichen Verlauf. Den Einfluss des Haber-Bosch-Prozesses sieht man darin, dass 50 Prozent des Stickstoffs im menschlichen Gewebe vom Haber-Bosch-Prozess stammen. Die Bevölkerungsexplosion und die Konzentration in den Städten wurden erst durch den Haber-Bosch-Prozess möglich, da er die Erzeugung von Nahrung für die große Bevölkerung und die Menschen in den Städten ermöglichte.

Die Hochskalierung begann 1909 durch Fritz Haber mit Hilfe von Hochdruckgeräten und Katalysatoren, um Ammoniak aus Luft herzustellen. Er bekam dafür 1918 den Nobelpreis. 1913 vergrößerte Carl Bosch, der bei BASF arbeitete, die Produktion durch eine großtechnische, kontinuierliche Hochdrucktechnik auf einen industriellen Maßstab. Er erhielt dafür 1931 den Nobelpreis. Also gab es Nobelpreise für die Hochskalierung und die Industrialisierung.

Phase 1: Grundlagenforschung

Die Grundlagenforschung bildet das Fundament, auf dem alle späteren Fortschritte und Anwendungen aufbauen. In dieser Phase geht es darum, die theoretischen und technologischen Grundlagen zu legen, die später eine breitere Nutzung ermöglichen. Historische Beispiele wie die Dampfmaschine oder der Haber-Bosch-Prozess verdeutlichen, wie entscheidend diese Phase für den langfristigen Erfolg ist.

Grundlagenforschung ist oft durch Unsicherheit und Experimentierfreude geprägt. Wissenschaftler und Ingenieure erkunden sehr viele neue Möglichkeiten und Ansätze, ohne genau zu wissen, welche Entdeckungen sich später als erfolgreich erweisen werden. Bei der Entwicklung der Dampfmaschine waren es beispielsweise erste Experimente mit der Umwandlung von Wasserdampf in mechanische Energie, die den Weg für die industrielle Nutzung ebneten. Beim Haber-Bosch-Prozess wiederum standen chemische Experimente im Vordergrund, bei denen es darum ging, Stickstoff aus der Luft zu binden und in Ammoniak umzuwandeln.

In der künstlichen Intelligenz begann die Grundlagenforschung bereits in den 1950er-Jahren. Die Pioniere dieser Zeit, wie Alan Turing, Marvin Minsky und John McCarthy, legten die theoretischen Grundlagen für die Entwicklung intelligenter Systeme. Sie stellten sich die Frage, wie Maschinen lernen, schlussfolgern und Probleme lösen können. Diese Ideen

waren revolutionär, doch ihre praktische Umsetzung war damals noch weit entfernt. Sie wurden während der ersten künstlichen Intelligenz-Konferenz 1956 in Dartmouth entwickelt. Dort prägte John McCarthy den Begriff »Artificial Intelligence«, künstliche Intelligenz.

Grundlagenforschung der KI

Die ersten Jahrzehnte der KI-Forschung waren bestimmt durch die Entwicklung von symbolischen Methoden wie logisches Erschließen, formale Methoden, Expertensysteme, Suchalgorithmen und Wissensrepräsentation.

Die subsymbolische KI hatte Probleme, sich zu etablieren, und erlebte ab 1969 ein großes Tief. Das 1958 von Frank Rosenblatt vorgestellte Perceptron war ein Lernalgorithmus für ein kleines Netzwerk ohne versteckte Schichten. Viele waren optimistisch, was subsymbolische KI mit neuronalen Netzen betraf. Dann veröffentlichten Marvin Minsky und Seymour Papert 1969 das Buch *Perceptrons: An Introduction to Computational Geometry*. Sie konnten zeigen, dass das Perceptron nicht mal das XOR-Problem lösen kann – ein sehr einfaches Problem aus der Informatik.

Hier setzte der KI-Winter ein, da viele Forscher enttäuscht von den Möglichkeiten der neuronalen Netze waren und es keine Förderungen für diese Forschungen mehr gab. In dieser Phase waren die Erwartungen oft überhöht, was zu Enttäuschungen führte. Der KI-Winter erstreckte sich über die 1970er- und 1980er-Jahre, in denen die Finanzierung stark zurückging, da die Ergebnisse hinter den Erwartungen zurückblieben. Dennoch legten diese Phasen den Grundstein für spätere Durchbrüche.

Das Ende des KI-Winters und die Wiederbelebung der Forschung an neuronalen Netzen begann 1986 mit den Arbeiten von David Rumelhart, Geoffrey Hinton und Ronald Williams, die den Backpropagation-Algorithmus für neuronale Netze wiederentdeckten. Vorläufer gab es schon

in den 60er-Jahren. Hier begann die Renaissance der neuronalen Netze, die bis Mitte der 1990er-Jahre anhielt. Ab Mitte der 1990er galten neuronale Netze im maschinellen Lernen als veraltet, da man sich auf Methoden wie »Support Vector Machines« oder »Kernel Methods« konzentrierte. Erst in den 2010er-Jahren kamen die neuronalen Netze als tiefe Netze wieder auf und wurden unter dem neuen Brandnamen »Deep Learning« populär. Einerseits wurde der Erfolg von Deep Learning durch die vielen Daten ermöglicht, die es nun gab, insbesondere durch das Internet. Andererseits wurde Deep Learning durch den Fortschritt in der Hardware erfolgreich, insbesondere durch die Entwicklung leistungsstarker GPUs, die das Lernen großer Modelle erlaubten.

Bezüglich der Grundlagenforschung zu Deep Learning gab es Vorarbeiten. 1965 trainierten Oleksij Iwachnenko und Walentyn Lapa tiefe Netze Schicht für Schicht. 1967 trainierte Shun-Ichi Amari mehrschichtige neuronale Netze durch stochastischen Gradientenabstieg (5 Schichten). 1970 veröffentlichte Seppo Linnainmaa einen Algorithmus für das Lernen in Netzen mit differenzierbaren Knotenpunkten. 1982 schlug Paul Werbos vor, die Methode für das Training neuronaler Netze zu verwenden.

Bezüglich der Grundlagenforschung zu »convolutional« neuronalen Netzen, also Netzen, die besonders gut für die Verarbeitung von Bild- und Videodaten geeignet sind, gab es ebenfalls Vorarbeiten. 1979 schlug Kunihiko Fukushima das Neocognitron vor: ein prinzipiell convolutional neuronales Netz. 1987 schlug Alex Waibel »time-delay« neuronale Netze mit Convolution, Gewichtsteilung und Backpropagation vor. Seit 1989 haben Yann LeCun und sein Team zahlreiche Verbesserungen an convolutional neuronalen Netzen gemacht. Yann LeCun ist durch seine Beiträge zu convolutional neuronalen Netzen sehr bekannt geworden und hat dafür den Turing Award bekommen.

Es gab auch Grundlagenforschung zu rekurrenten neuronalen Netzen. Im Jahr 1920 stellten Ernst Ising und Wilhelm Lenz die erste nicht lernende rekurrente neuronale Netzarchitektur vor. Im Jahr 1943 befassten sich Warren McCulloch und Walter Pitts mit rekurrenten neuronalen Netzen.

1956 analysierte Stephen Cole Kleene rekurrente neuronale Netze. 1972 veröffentlichte Shun-Ichi Amari lernende rekurrente neuronale Netze. 1986 schlug Mike Jordan eine rekurrente neuronale Netzarchitektur vor, das »Jordannet«. 1991 entwickelte ich in meiner Diplomarbeit das Long Short-Term Memory (LSTM).

Ohne die jahrzehntelange Grundlagenforschung wären die heutigen Erfolge der KI nicht möglich gewesen. Textmodelle wie ChatGPT, generative Bildverarbeitungssysteme und die Fortschritte beim autonomen Fahren basieren auf den theoretischen und technologischen Fortschritten, die in dieser Phase der Grundlagenforschung erzielt wurden. Doch die Grundlagenforschung endet nie wirklich: Sie bleibt ein kontinuierlicher Prozess, der neue Möglichkeiten erkundet und die Basis für zukünftige Innovationen legt.

Phase 2: Skalierung

Mit der Grundlagenforschung als Fundament begann sich die moderne KI ab 2009 von einer rein theoretischen Disziplin zu einer praxisrelevanten Technologie zu entwickeln. Ab 2009 hatte die moderne KI große Erfolge bei Wettbewerben, in denen sich die besten Methoden in der Sprachverarbeitung und Bildverarbeitung verglichen. Man denke nur an den ImageNet-Moment durch AlexNet. Ab 2017 begann dann die nächste Phase, die Skalierung, mit der ResNet-Architektur und der Transformer-Architektur. Sie markiert den Übergang von experimentellen Konzepten zu groß angelegten Anwendungen.

Während bei der Grundlagenforschung viele Methoden getestet werden, wird die Methodenvielfalt bei der Skalierung viel kleiner. Nur wenige Methoden eignen sich zur Hochskalierung, wie wir sowohl bei der Dampfmaschine als auch dem Haber-Bosch-Prozess schon gesehen haben. Bei der Dampfmaschine war es die Konstruktion mit Kolben und Fliehkraftregler zur Geschwindigkeitsregulierung, die die Hochskalierung

ermöglichte. Beim Haber-Bosch-Prozess waren effiziente Katalysatoren und der hohe Druck entscheidend für das Hochskalieren. In beiden Fällen wurde die Methodenvielzahl stark eingeschränkt.

Genau das sehen wir jetzt auch in der modernen KI, die sich auf die ResNet- und die Transformer-Architektur konzentriert. Diese Architekturen lassen eine Hochskalierung der Netze, also Netze mit Milliarden von Parametern und Hunderten von Neuronenschichten, zu. Das ist frustrierend für KI-Grundlagenforscher, da die meisten neuen Methoden für die Hochskalierung nicht geeignet sind oder nicht getestet werden können, da dies zu teuer wäre. Diesen Frust artikulierte der US-Informatiker und Turing-Award-Gewinner Richard Sutton, einer der Väter des Reinforcement Learning, in seinem Artikel *The Bitter Lesson*, in dem er bedauert, dass Fortschritte in der KI nur durch Hochskalierung gemacht wurden.

In dieser Skalierungsphase stehen große Datenmengen, Hardware-Ressourcen, Optimierungen für die Hardware, Datenaufbereitung und das Orchestrieren von großen Rechenclustern im Vordergrund. Historische Technologien wie die Dampfmaschine und der Haber-Bosch-Prozess zeigen, wie entscheidend die Skalierung für den wirtschaftlichen Durchbruch und die gesellschaftliche Verbreitung einer Erfindung ist.

In dieser Phase der Skalierung geht es darum, die Technologie »groß zu machen«. Bei der Dampfmaschine bedeutete dies, größere Maschinen mit höherem Druck und effizienterer Leistung zu bauen. Beim Haber-Bosch-Prozess führte die Skalierung zu einer massiven Steigerung der Produktionskapazität, was schließlich, wie schon oben erwähnt, zur Verleihung von Nobelpreisen führte – einmal für die Hochskalierung (Haber) und einmal für die Industrialisierung (Bosch bei BASF).

Für die KI bedeutet diese Phase, die Modelle, die Trainingsdaten und die Rechenkapazitäten dramatisch zu vergrößern. Die Veröffentlichung von Textgenerierungsmodellen wie ChatGPT oder Bildgenerierungsmodellen wie DALL-E, Midjourney und Stable Diffusion zeigt, dass wir mitten in dieser Phase stecken. Hier ist allerdings nur Stable Diffusion ein Open-Source-Projekt, was, wie oben beschrieben, wünschenswert für die

Forschung wäre. Große Text- und Bildgenerierungsmodelle beeindrucken durch ihre Leistungsfähigkeit. Large Language Models (LLMs) können Texte erzeugen und Diffusionsmodelle können Bilder auf der Basis eines Prompts generieren, sodass man die generierten Texte und Bilder von solchen, die von Menschen gemacht wurden, nicht mehr unterscheiden kann. Doch diese Hochskalierungsphase ist auch geprägt von enormem Ressourcenbedarf hinsichtlich Daten und Rechenleistung. Das macht ihre Konstruktion oft teuer und ineffizient. Aber auch bei der Inferenz, das heißt bei ihrer Anwendung, sind die Modelle teuer, da sie viel Rechenleistung benötigen.

Nach der Skalierungsphase startet die Suche nach Anwendungsmöglichkeiten. Die hochskalierten Modelle liefern beeindruckende Leistungen, aber wie und wofür soll man sie in der Praxis verwenden? Unternehmen und Forschungseinrichtungen experimentieren mit den neuen Technologien, doch die Integration in den Alltag bleibt herausfordernd. Viele Firmen wissen nicht genau, wie sie diese leistungsstarken Modelle produktiv einsetzen sollen. Das ist vergleichbar mit den frühen Tagen der Elektrizität: Man wusste, dass die Technologie revolutionär war, aber die konkrete Umsetzung in Haushalten und Fabriken brauchte Zeit.

Während in der Grundlagenforschung der Dampfmaschine die Experimente mit Dampf und Druck im Fokus standen, ermöglichte die Skalierung den Bau größerer, effizienterer Maschinen, die in Fabriken, der Logistik und der Mobilität eingesetzt wurden. Der Haber-Bosch-Prozess, ursprünglich eine chemische Innovation, wurde durch die Skalierung zu einem Verfahren, das in großtechnischen Anlagen die globale Landwirtschaft revolutionierte, indem Dünger in großen Mengen produziert werden konnte.

Bei der künstlichen Intelligenz zeigt sich die Skalierung vor allem in der Entwicklung großer Modelle wie GPT-1–4/ChatGPT (OpenAI), Gemma/Gemini/mT5 (Google), Llama sowie OPT (Meta), Megatron (Nvidia), OLMo (Allen Institute), DeepSeek-v3 (DeepSeek), Mistral

(Mistral) oder Claude (Anthropic), die durch enorme Rechenleistung und an umfangreichen Datenkorpora trainiert werden. Während die Grundlagenforschung die Basis für die mathematischen Optimierungen und architektonischen Durchbrüche legte, steht in der Skalierungsphase die Implementierung in großem Maßstab im Vordergrund. Diese Phase ist geprägt von einem intensiven Ressourceneinsatz: Gigantische Datenmengen müssen gesammelt, vorverarbeitet, selektiert und strukturiert werden. Gleichzeitig ist der Bedarf an Rechenleistung enorm, was nur durch den Einsatz spezialisierter Hardware wie Graphics Processing Units (GPUs) oder Tensor Processing Units (TPUs) zu bewerkstelligen ist.

Ein prägnantes Beispiel für die Skalierung in der KI ist die Entwicklung generativer Modelle zur Texterzeugung. Systeme wie GPT-4 haben bewiesen, dass die Vergrößerung der Modellarchitektur – etwa durch mehr Parameter und mehr Schichten – die Leistung erheblich verbessern kann. Diese Modelle sind in der Lage, Texte mit einer Präzision und Vielseitigkeit zu generieren, die in der Phase der Grundlagenforschung noch undenkbar war. Nach der Skalierung versucht man, diese Modelle in verschiedensten Anwendungen einzusetzen, von der automatisierten Kundenkommunikation bis hin zur Unterstützung wissenschaftlicher Forschung. Jedoch bleibt bei den Anwendungen noch viel zu tun.

Neben generativen Textmodellen sind auch generative Bildmodelle wie Midjourney Beispiele für die Skalierung von KI. Sie demonstrieren, wie durch massive Datenmengen und Rechenleistung konstruktive Prozesse automatisiert und optimiert werden können. In der Praxis wird dies bereits in der Werbebranche oder im Design genutzt, wo KI dabei hilft, visuelle Inhalte effizient zu erstellen.

Der hohe Bedarf an Rechenressourcen und die Fokussierung auf datenintensive Anwendungen werfen Fragen nach der Energieeffizienz und den ökologischen Auswirkungen auf. Zudem bleibt die Skalierung nicht allen zugänglich: Die Entwicklung und der Betrieb großer KI-Modelle sind extrem kostspielig, was dazu führt, dass diese Technologien oft in den Händen weniger Großkonzerne konzentriert bleiben. Dies birgt das Risiko,

dass kleinere Akteure vom technologischen Fortschritt ausgeschlossen werden. Hier sollte es zu einer Demokratisierung der KI kommen, die noch im Besitz der großen Unternehmen aus den USA und China ist. Open Source wäre ein Ansatz, der den Weg weisen könnte.

Zusammenfassend kann man die Skalierungsphase als Phase der Etablierung einer Technologie beschreiben. Sie hebt die Innovation aus dem Labor in die reale Welt und demonstriert ihren möglichen Nutzen, jedoch erst nur in Spezialanwendungen wie Textgenerierung und Bildgenerierung. Diese Phase ist entscheidend, um die Grundlagen für die Industrialisierung zu legen, in der KI noch in vielen Domänen, Anwendungen und Gebieten eingesetzt wird. Die Skalierung zeigt, wie eine Technologie durch ihren Einsatz in großem Maßstab eingesetzt werden könnte, um die Produktivität in der Wirtschaft zu steigern.

Phase 3: Industrialisierung

Die dritte Phase ist die Industrialisierung, in der eine Technologie in den Kern wirtschaftlicher Prozesse integriert wird. Die Dampfmaschine revolutionierte nicht nur die Textilproduktion, sondern legte auch den Grundstein für neue Transportmöglichkeiten wie Eisenbahnen und Dampfschiffe. Ebenso brachte der Haber-Bosch-Prozess die industrielle Ammoniaksynthese hervor, die die landwirtschaftliche Produktivität durch massenhaft verfügbaren Dünger dramatisch steigerte.

In der Phase der Industrialisierung ist bei der Dampfmaschine und beim Haber-Bosch-Prozess die Methodenvielfalt wieder gewachsen. Die verschiedenen Branchen haben verschiedene Bedürfnisse, Einschränkungen und Randbedingungen, was Modifikationen der Basismethode notwendig macht. Hier gibt es wieder viele Möglichkeiten, wie man die bestehenden hochskalierten Methoden auf die Bedürfnisse der speziellen Anwendung anpasst. An dieser Stelle ist also wieder die Vielfalt der Forschung gefragt, neue Ideen und Innovationen sind notwendig.

In der KI-Entwicklung stehen wir gerade an der Schwelle zu dieser Phase. Aktuell sind die meisten Anwendungen noch sehr speziell, experimentell oder ergänzen bestehende Prozesse lediglich am Rande. Doch die Industrialisierung der KI wird kommen. Sie wird kleinere, effizientere Modelle umfassen, die spezifische Aufgaben lösen und direkt in Produktions- und Entwicklungsprozesse eingebunden werden. Der Frust der KI-Forscher ist dann beseitigt, da auch in der KI Ideen und Innovation wieder gefragt sind und »The Bitter Lesson« vorbei ist.

Ein Beispiel für die Industrialisierung der KI ist die Anwendung von KI in der Fertigung. Anstatt lediglich unterstützende Funktionen wie Qualitätskontrollen auszuführen, könnten KI-Systeme direkt in den Designprozess eingreifen und neue, optimierte Produkte entwickeln. Diese Integration wird nicht nur Effizienzgewinne bringen, sondern auch die Innovationsgeschwindigkeit drastisch erhöhen.

Die Industrialisierung stellt den entscheidenden Schritt dar, in dem neue Technologien aus der experimentellen Forschung und den ersten Anwendungen in die breite und systematische Nutzung übergehen. Historische Beispiele wie die Dampfmaschine oder der Haber-Bosch-Prozess verdeutlichen, wie technologische Innovationen die Wirtschaft und Gesellschaft nachhaltig verändern können.

Die Industrialisierung der KI bedeutet, dass diese Technologie nicht länger nur in Forschungslabors oder als Pilotprojekt eingesetzt, sondern zu einem festen Bestandteil der Kernprozesse in Wirtschaft und Industrie wird. Während die Skalierungsphase von universellen und breit einsetzbaren Modellen geprägt war, rückt in der Industrialisierungsphase die Entwicklung spezialisierter KI-Lösungen in den Vordergrund.

Simulation

Einen besonderen Stellenwert nimmt in der Industrialisierungsphase die Simulation ein. In unserem Kontext sind Simulationen digitale

Nachbildungen von realen Prozessen oder Systemen, die auf physikalischen Gesetzen basieren. Simulation ist immer ein dynamischer Prozess in der Zeit, das heißt, ein Prozess wird in der Zeit vorangetrieben. Die physikalische Welt wird am Computer imitiert. Simulationen ermöglichen es, physikalische, chemische oder mechanische Prozesse am Computer zu modellieren und zu optimieren, bevor physische Experimente oder Tests durchgeführt werden. Traditionelle numerische Simulationen stoßen jedoch bei der Abbildung komplexer Systeme an ihre Grenzen aufgrund der extrem hohen Anzahl von Teilchen oder Gitterpunkten. KI-gestützte Simulationen bieten hier einen entscheidenden Vorteil: Sie erkennen Muster und können dadurch viele Teilchen oder Gitterpunkte durch ein Muster ersetzen, wodurch die Rechenzeit drastisch reduziert wird. Beispielsweise könnten alle Teilchen einer Flüssigkeit, die an einem Strudel beteiligt sind, durch den Strudel an sich ersetzt werden. Wir beschreiben auch nicht alle Teilchen unseres Mondes, sondern fassen diese durch Ort, Geschwindigkeit plus Richtung und Masse des Mondes zusammen. Der Mond wird durch ein Objekt dargestellt, da sich alle seine Teilchen synchron bewegen.

Numerische Simulationen bilden komplexe physikalische Prozesse ab, wie sie in Hochöfen, Strömungsmechaniken (Flugzeug oder Auto) oder chemischen Reaktoren vorkommen. Diese Simulationen sind präzise, doch die Anzahl der Teilchen, Gitterpunkte oder Parameter, die simuliert werden müssen, steigt mit der Komplexität des Systems. Komplexere Systeme brauchen eine genauere Auflösung sowohl im Raum als auch in der Zeit. Für industrielle Prozesse, die Hunderte Millionen von Teilchen betreffen, bedeutet dies oft, dass nur kleine Bereiche oder kurze Zeiträume simuliert werden können. Simulationen dieser Art können Wochen oder Monate dauern – ein Luxus, den viele Branchen nicht haben.

Ein Beispiel ist die Simulation von Hitzeverteilungen in Hochöfen: Solche Systeme bestehen aus unzähligen Gitterpunkten, deren Verhalten genau vorhergesagt werden muss, um optimale Ergebnisse zu erzielen. Die klassische numerische Simulation kann jedoch nur ein grobes Netz aus Datenpunkten liefern, das nicht alle Details abdeckt. Eine Verfeinerung des

Modells ist technisch möglich, erfordert aber enorme Rechenressourcen und somit Zeit.

Hier kommt KI ins Spiel. Anders als klassische Methoden ist KI in der Lage, Muster zu erkennen und komplexe Systeme in abstrahierte Modelle zu übersetzen. Diese Modelle besitzen sogenannte »Hidden Spaces«, die die Datenmenge drastisch reduzieren, indem sie Muster erkennen und nur die wichtigsten Aspekte eines Prozesses abbilden. Für den Hochofen könnte dies bedeuten, dass nicht mehr jeder Gitterpunkt simuliert wird, sondern Gruppen von Gitterpunkten zu Makrostrukturen zusammengefasst werden. Die Interaktion dieser Makrostrukturen wird dann von der KI simuliert, wodurch sich die Rechenzeit erheblich verkürzt.

In der Automobilindustrie beispielsweise könnten KI-gestützte Simulationen dazu genutzt werden, um Oberflächenströmungen vorherzusagen und dadurch den Widerstand zu berechnen. Oder es werden die mechanischen Belastungen an Bauteilen präzise vorhergesagt, ohne dass umfangreiche physische Prototypen gebaut werden müssen. Dies beschleunigt nicht nur die Entwicklungsprozesse, sondern reduziert auch die Produktionskosten erheblich. In der chemischen Industrie können solche Modelle molekulare Wechselwirkungen in Reaktoren simulieren, um die Effizienz chemischer Reaktionen zu steigern und den Energieverbrauch zu senken.

An unserem Institut an der Johannes Kepler Universität Linz war das Interesse an Simulationen schon lange da, aber ein entscheidender Moment war, als Johannes Brandstetter nach seiner Zeit an der Universität Amsterdam und bei Microsoft zu uns zurückkam. Er arbeitete mit Max Welling, einem der führenden KI-Forscher in Europa, an großen Projekten, darunter das Wettermodell Aurora von Microsoft. Johannes brachte dieses Wissen zu uns, und wir begannen, tiefer in die Welt der Simulationen einzutauchen. Microsoft war zunächst sehr stark in Richtung Simulation orientiert, bevor sie sich dann doch auf Modelle zur Textgenerierung konzentrierten. Microsoft erkannte, dass Modelle zur Textgenerierung für ihre Geschäftsstrategie, insbesondere in Bezug auf das Internet und soziale

Medien, wichtiger waren. Sie hatten ursprünglich viel in Simulationsprojekte wie AI for Physics investiert, aber als sie merkten, dass Textmodelle wie GPT größere unmittelbare kommerzielle Relevanz haben, stellten sie viele dieser Simulationsprojekte ein. Dadurch entstand eine Lücke in der Simulation, die jetzt genutzt werden könnte.

Während die großen Tech-Firmen in den USA ihren Fokus auf Modelle zur Textgenerierung richten, haben wir in Europa die Chance, uns stärker auf die Simulationen zu konzentrieren. Viele europäische Unternehmen sind in der produzierenden Industrie oder im Maschinenbau tätig, wo Simulationen eine viel größere Rolle spielen. Während Modelle zur Textgenerierung für Internetfirmen entscheidend sind, ist die Simulation für unsere Industrien in Europa wesentlich relevanter. Das ist der Bereich, in den wir investieren sollten.

Die Industrialisierung von KI-gestützten Simulationen bringt zahlreiche Vorteile mit sich. Prozesse können schneller und kosteneffizienter entwickelt werden, da physische Prototypen weitgehend entfallen. Durch die hohe Präzision der KI-Modelle lassen sich Fehler bereits im virtuellen Raum identifizieren und beheben. Dies führt nicht nur zu Kosteneinsparungen, sondern auch zu einem nachhaltigeren Umgang mit Ressourcen, da Material- und Energieverbrauch reduziert werden. Die Simulation wird somit zu einem zentralen Werkzeug, um Innovationsprozesse zu beschleunigen und die Produktivität in der Industrie zu steigern.

Im Klimaschutz bieten KI-gestützte Simulationen enorme Chancen: Sie könnten dazu beitragen, industrielle Prozesse klimafreundlicher zu gestalten, den Energieverbrauch zu senken oder CO_2-Emissionen präzise zu reduzieren. Der Erfolg von LSTM in der Hydrologie wurde bereits erwähnt. Hier wird der Wasserstand von Flüssen simuliert, wodurch man Fluten vorhersagen kann. Man kann aber auch vorhersagen, was geschieht, wenn sich die Temperatur um wenige Grad erhöht. Solche Simulationen sind wichtig, um die Auswirkungen des Klimawandels zu modellieren.

Während traditionelle Simulationsmethoden bereits in Bereichen wie der Automobilindustrie, der Medizin und der Chemie etabliert sind, stoßen sie bei zunehmender Komplexität der Systeme an ihre Grenzen. KI-basierte Simulationen erweitern die Möglichkeiten aber erheblich, indem sie Muster und Makrostrukturen erkennen und somit Vorhersagen viel schneller machen können.

6

Die Welt simulieren

Simulationen sind digitale Nachbildungen von realen Prozessen oder Systemen, die über die Zeit laufen und die auf physikalischen Gesetzen basieren. Darin unterscheiden sie sich von der Modellierung, die nicht auf physikalische Prozesse beschränkt ist oder nicht über die Zeit laufen muss. Eine Pandemie etwa kann man modellieren, aber nicht simulieren, da die physikalischen Gesetze nicht vorhanden sind. Simulieren kann man hingegen Gas- oder Flüssigkeitsflüsse, die mit partiellen Differentialgleichungen beschrieben werden können.

Die Simulation ermöglicht es, Modelle großer physikalischer Systeme zu bauen, Zustände vorherzusagen, komplexe Zusammenhänge zu verstehen, Hypothesen zu testen und Entscheidungen zu optimieren – ohne die Kosten und Risiken, die physische Experimente oft mit sich bringen. Ob es um die Entwicklung von neuen Materialien, die Optimierung industrieller physikalischer oder chemischer Produktionsprozesse, die Optimierung von Wasserturbinen oder Lufträdern, die Optimierung des Luftwiderstandes von Fahrzeugen, das Verhalten im Inneren von Hochöfen, das Verhalten von Teilchen in großen Fülltrichtern oder das Plasmaverhalten bei Kernfusion geht: KI-gestützte Simulationen könnten die Grundlage für nachhaltige Innovationen und damit für die nächste industrielle Revolution bilden.

Historisch betrachtet, waren Simulationen schon immer ein wichtiger Bestandteil des wissenschaftlichen und industriellen Fortschritts. Ingenieure nutzten sie, um die Aerodynamik von Fahrzeugen zu verbessern, Chemiker, um komplexe Reaktionen zu verstehen, und Meteorologen, um das Wetter vorherzusagen. Was jedoch neu ist, ist die Geschwindigkeit, Präzision und Flexibilität, die KI diesen Prozessen verleiht. Mit ihrer Hilfe können

Simulationen nun in Echtzeit durchgeführt werden, wobei Millionen von Variablen berücksichtigt und optimiert werden können. Diese Entwicklung markiert einen Paradigmenwechsel: von der isolierten Anwendung einzelner Simulationen hin zu einem integrativen Ansatz, bei dem KI als zentraler Motor für Innovation und Effizienz fungiert.

Eine konkrete Anwendung ist die Optimierung von Produktionsprozessen, etwa in der Pharmaindustrie, wo Tabletten aus einer genauen Mischung von Bestandteilen bestehen sollten. Sie sollten alle gleich viel Wirksubstanz haben oder mit Schichten von vorgegebener Dicke überzogen sein. Diese Prozesse müssen optimiert werden, um effizienter zu produzieren. Mit herkömmlichen Methoden würde es Monate dauern, diese Prozesse im industriellen Maßstab zu simulieren. Die KI kann jedoch Muster, wie häufig vorkommende lokale Vermischungsmuster, lernen und dadurch die Simulation erheblich beschleunigen.

Ein weiteres Beispiel ist die Energieerzeugung. Durch Simulationen könnte man herausfinden, wie man den Strom effizienter durch Leitungen transportiert oder in Transformatoren umwandelt, um Verluste durch Wärmeabgabe zu minimieren und den Durchsatz zu maximieren. In beiden Fällen würde die KI die Geschwindigkeit der Simulationen im Vergleich zu numerischen Verfahren erhöhen.

Insgesamt können Simulationen unsere Welt vorhersagbarer und beherrschbarer machen. Denn sie ermöglichen das Durchspielen von physikalischen Prozessen in einem sicheren und kontrollierten Rahmen. Das ist besonders nützlich, wenn physische Experimente zu kostspielig, zeitaufwendig oder riskant wären. Mit der Unterstützung von KI können Simulationen dynamischer gestaltet werden. In der Simulation kann man Systeme an die Grenzen treiben, ohne dass für Menschen oder Anlagen Gefahr besteht, wie beispielsweise bei der Simulation der Kernfusion.

Die Beschleunigung von Simulationen durch KI basiert auf der Annahme, dass KI Muster (Abstraktionen und Makrostrukturen) identifizieren kann. Diese Muster erlauben es, komplexe Prozesse auf ein handhabbares Maß zu reduzieren, ohne die wesentlichen Eigenschaften des

Systems zu verlieren. Nehmen wir die Simulation von Luftströmungen um ein Fahrzeug: Statt jedes einzelne Luftmolekül zu simulieren, werden große aerodynamische Bewegungsmuster erkannt, die dann als Makrostrukturen simuliert werden und damit die Simulation erheblich beschleunigen.

Die KI ist in der Lage, versteckte Muster und Beziehungen in den Daten zu erkennen, die für die menschliche Intuition oft unsichtbar bleiben. Dadurch können Simulationen schneller und genauer gestaltet werden, was wiederum neue Möglichkeiten in der Forschung und Industrie eröffnet.

Die Macht der Abstraktion

Abstraktion ist der Schlüssel zur Effizienz: Anstatt jedes Detail eines Systems auf atomarer Ebene zu simulieren, konzentriert man sich auf die Aspekte, die für das jeweilige Ziel relevant sind. Der große Vorteil dieser abstrakten Modelle ist ihre Fähigkeit, Zeit und Ressourcen zu sparen.

Traditionelle Simulationsmethoden, die numerische physikalische Modelle nutzen, arbeiten oft mit einem direkten Ansatz: Jeder Teil des Systems wird auf Basis von physikalischen oder chemischen Gesetzen einzeln betrachtet und berechnet. Leider kennen wir die Physik nur für kleinste Teilchen wie Moleküle. Wir wissen nicht, wie Sandkörner interagieren, wie Getreidekörner aufeinander Einfluss nehmen, wie Maiskörner aufeinander wirken oder wie Tabletten sich gegenseitig beeinflussen. Die Simulation auf physikalischer Ebene, das heißt auf Molekülebene, ist nicht nur zeit- und ressourcenintensiv, sondern stößt bei hochkomplexen Systemen schnell an ihre Grenzen. KI hingegen kann lernen, welche Details für die Simulation wirklich relevant sind, und den Fokus auf die wesentlichen Aspekte und Makrostrukturen legen. Diese Fähigkeit der Abstraktion ermöglicht es ihr, die Komplexität drastisch zu reduzieren, ohne an Genauigkeit einzubüßen. So kann KI die Wirkung von Sand- oder Getreidekörnern aufeinander lernen – also die Physik für Sand- oder Getreidekörner. Nun können ganze Körner simuliert werden und nicht

mehr nur deren Moleküle. Die Körner bilden also die Makrostrukturen, die dann als ein Objekt simuliert werden können.

Ein gutes Beispiel für die Macht der Abstraktion ist die Physik. Das Gravitationsgesetz abstrahiert die Bewegung von Planeten, Äpfeln oder anderen Objekten und fasst sie in eine einfache Formel. Diese Abstraktion hat es uns ermöglicht, das Universum zu verstehen, ohne jedes einzelne Objekt, zum Beispiel jedes Molekül, darin zu analysieren.

Ebenso abstrahiert Sprache die Welt, um Wissen effizient zu vermitteln. Wenn wir sagen, dass ein Objekt ein »Laubbaum« ist, abstrahieren wir viele Details – die Anzahl und Form der Blätter, die Farbe des Stamms, die genaue Form der Äste. Diese Abstraktion macht es uns leichter, die Welt zu verstehen und zu kommunizieren. Simulationen nutzen denselben Ansatz: Sie abstrahieren, um die Essenz eines Systems zu erfassen, ohne jedes Detail nachzubilden.

Sprache ist eines der mächtigsten Werkzeuge der Menschheit, um die Welt zu abstrahieren. Wenn wir sagen, dass etwas ein »Tiger« ist, abstrahieren wir von den individuellen Merkmalen eines spezifischen Tigers. Dieses Prinzip erlaubt es uns, Wissen weiterzugeben – sei es über gefährliche Tiere, essbare oder giftige Pflanzen oder nützliche Werkzeuge. Die Fähigkeit zur Abstraktion war daher ein entscheidender Schritt in der Entwicklung der menschlichen Zivilisation.

Simulationen funktionieren auf ähnliche Weise. Sie abstrahieren die Realität, um allgemeine Regeln und Strukturen zu finden, die auf viele verschiedene Situationen anwendbar sind.

Stellen Sie sich vor, wir möchten die Bewegung eines Stuhls modellieren. Anstatt jedes einzelne Atom des Stuhls und der Umgebung zu simulieren, abstrahieren wir. Wir beschreiben den Stuhl als eine Einheit, definieren seine Masse, die Reibung der Beine auf dem Boden und berücksichtigen vielleicht 20 weitere Parameter. Auf diese Weise können wir das Verhalten des Stuhls ausreichend genau modellieren, ohne in die kleinsten Details zu gehen.

Die Grundprinzipien sind bei Simulationen und Textmodellen sehr ähnlich. Textmodelle wie GPT lernen, Muster in Texten zu erkennen und Vorhersagen über das nächste Wort zu treffen. Bei Simulationen hingegen geht es darum, Muster im aktuellen Zustand und den vergangenen Zuständen zu finden und den nächsten Zustand präzise vorherzusagen.

Während Textmodelle sich auf Sequenzen von Wörtern oder Tokens konzentrieren, arbeiten Simulationsmodelle oft mit räumlichen Daten, wie etwa Gitterpunkten in einem Hochofen. Datenpunkte in der Simulation könnten daher Temperatur, Druck oder andere physikalische Eigenschaften an einem bestimmten Ort beschreiben. Diese Datenpunkte sind jedoch nicht isoliert: Sie stehen in komplexen räumlichen und zeitlichen Beziehungen zueinander, denn die Materie an den Gitterpunkten unterliegt den physikalischen Gesetzen, die das Modell verstehen muss.

Transformer-Modelle, die für Sprachverarbeitung entwickelt wurden, können auch in der Simulation eingesetzt werden. Sie können sowohl räumliche Muster in einem Gitter oder einer Menge von Partikeln erkennen als auch zeitliche Entwicklungen prognostizieren. Allerdings ist dies nur eine von vielen möglichen Herangehensweisen. Simulationen erfordern oft maßgeschneiderte Modelle, die speziell für die jeweilige Domäne entwickelt werden.

Die größte Herausforderung bei der KI-gestützten Simulation liegt in der Datenkomplexität. Klassische Simulationen erfordern oft extrem kleine Zeitschritte und hohe räumliche Auflösungen, um die physikalischen Prozesse akkurat abzubilden. Die KI muss lernen, diese Komplexität zu reduzieren, ohne an Genauigkeit zu verlieren.

Für Simulationen sind oft spezialisierte Modelle gefragt, die sowohl mit großen Datenmengen umgehen können als auch in der Lage sind, räumliche und zeitliche Zusammenhänge zu erkennen. Das erfordert eine Kombination aus Deep Learning, Optimierungsmethoden und domänenspezifischem Wissen. Die Zukunft der KI-Simulationen liegt in ihrer Fähigkeit, immer präzisere und universellere Abstraktionen zu schaffen. Dies erfordert technologische Fortschritte.

Physik lernen

Die Frage, die mich am meisten fasziniert, ist: Wie können wir die Welt um uns herum besser verstehen? Für mich liegt der Schlüssel in der Entwicklung von Simulationen, die die Sprache der Wirklichkeit sprechen – je abstrakter diese Sprache ist, desto besser wird die reale Welt verstanden und desto schneller kann man simulieren. Die KI muss lernen, nicht jedes Detail exakt nachzubilden, sondern die zugrunde liegenden Muster und Gesetze zu erkennen. Dasselbe gilt für die Modellierung komplexerer Systeme wie das Wetter, die Zelle oder sogar eine ganze Stadt. Wenn wir die »Sprache« eines Systems verstehen, können wir Vorhersagen treffen und Probleme lösen, ohne jedes einzelne Detail zu modellieren oder zu simulieren.

Bei der Modellierung von Systemen, die für Menschen zu komplex sind, kommt die KI ins Spiel. Eine KI kann komplizierte und vielschichtige Muster erkennen, die für uns einfach nicht zu verstehen sind. Sie kann Beziehungen zwischen so vielen Variablen herstellen, die wir aufgrund ihrer zu großen Menge nicht erfassen könnten.

Ein Beispiel ist eine Modellierung im Bereich des Klimas: Wenn wir verstehen, wie Meeresströmungen, Luftverschmutzung und Wetterphänomene zusammenwirken, können wir bessere Maßnahmen gegen den Klimawandel ergreifen. Die einzelnen physikalischen Phänomene, wie etwa die Meeresströmungen, können simuliert und dann in ein globales Modell integriert werden.

Simulationen nutzen Abstraktion, um allgemeine Regeln zu finden und auf verschiedene Szenarien anzuwenden. Ein gutes Beispiel ist die Simulation chemischer Reaktionen: Während traditionelle Ansätze Molekülwechselwirkungen einzeln berechnen, analysiert KI Muster von Hitze-, Material- und Katalysatorverteilungen und erstellt damit schnell präzise Vorhersagen. Dieses Prinzip findet Anwendung in der

Materialforschung, der Produktionsoptimierung und der Medikamentenherstellung.

In der Materialforschung könnten KI-Modelle genutzt werden, um vorherzusagen, wie ein bestimmtes Material unter extremen Bedingungen reagiert, ohne dass es tatsächlich getestet werden muss. Ähnlich könnte in der Energiebranche die Effizienz von Wind- oder Wasserturbinen durch KI-gestützte Simulationen maximiert werden, indem Faktoren wie Wetterbedingungen oder Energieverbrauch in Echtzeit berücksichtigt werden. Insbesondere kann die Simulation in Zustandsbereiche gehen, wo ein teures Gerät oder die Anlage beschädigt werden oder sogar explodieren würde. Die Simulation birgt diese Gefahren für Mensch und Geräte nicht.

Während numerische Simulationen auf physikalischen Gesetzen basieren, lernen KI-Modelle die Gesetze aus den Daten. Diese gelernten Gesetze sind meist nicht so gut validiert wie unsere physikalischen Gesetze. Um dieses Problem zu lösen, arbeiten Wissenschaftler weltweit daran, Methoden zu entwickeln, die die Ergebnisse von KI-gestützten Simulationen validieren und nachvollziehbar machen. Die Welt zu verstehen, bedeutet oft, die Komplexität zu reduzieren. Wir abstrahieren, gruppieren und ordnen – sei es in der Physik, der Biologie oder der Sprache. Diesem Prinzip folgt auch die künstliche Intelligenz.

Im Maschinen- und Anlagenbau wird oft die Simulation von Luftströmen oder Flüssigkeiten benötigt, um beispielsweise den Luftwiderstand von Autos oder Flugzeugen zu reduzieren oder Heizungen, Turbinen und Pumpen zu optimieren. Aber auch natürliche Strömungen wie bei Flüssen oder im Meer müssen simuliert werden, wenn man die Auswirkungen von Eingriffen wissen möchte. Für solche Simulationen könnten KI-Systeme genutzt werden, um Muster wie Wirbel oder Strömungsmuster zu identifizieren und als Einheit zu simulieren. Statt jedes einzelne Teilchen der Flüssigkeit oder des Gases zu simulieren, können wir die erkannten Muster beschreiben und so komplexe Systeme effizienter simulieren, da diese Muster wie Wirbel als Einzelobjekte betrachtet werden.

Die Simulation ähnelt dem Planen, das auch stark auf einer Abstraktion und einem Weltmodell basiert. Beim Planen konzentrieren wir uns auf die wesentlichen Faktoren einer Situation und blenden irrelevante Details aus. Wenn wir beispielsweise einen Flug zum Flughafen planen, berücksichtigen wir Verkehrsstaus, Flugzeiten und Abfahrtszeiten öffentlicher Verkehrsmittel – die Farbe des Taxis spielt hingegen keine Rolle. Diese Fähigkeit zur Abstraktion ermöglicht es uns, komplexe Entscheidungen schnell zu treffen, ohne von den Details überwältigt zu werden. Hier verwenden wir keine Simulation, sondern ein Weltmodell. Planen hat im Gegensatz zur Simulation ein konkretes Ergebnis vor Augen, sei es, pünktlich am Flughafen anzukommen oder ein Projekt erfolgreich abzuschließen.

Die Simulation imitiert zeitliche Prozesse der realen Welt, die exakt und schnell approximiert werden sollen. Mit Simulationen kann man Prozesse verbessern, die Umgebung an Prozesse anpassen, die Prozesse besser kontrollieren oder neue erfinden. Hingegen ist ein Weltmodell meist nicht sehr exakt und dient dazu, solche Aktionen auszuführen, die die Welt in gewünschter Weise beeinflussen oder dabei helfen, Ziele zu erreichen. Weltmodelle sind immer mit einem Agenten verbunden, der sie benutzt und die Umgebung beeinflusst. Weltmodelle helfen, vorherzusagen, was unsere Aktionen in der Welt auslösen und wie die Welt auf uns wirkt. Auch Mäuse oder Vögel haben ein Weltmodell, das aber sehr viel einfacher ist als das menschliche.

Die Grenzen des menschlichen Denkens werden sichtbar, wenn Systeme zu komplex werden. Ein gutes Beispiel hierfür sind meteorologische Vorhersagen: Während Menschen in der Lage sind, allgemeine Wettertrends zu verstehen, ist die präzise Vorhersage von Temperatur- und Niederschlagsmustern ohne den Einsatz von Simulationen und KI praktisch unmöglich. Hier zeigt sich die Stärke der KI: Ihre Fähigkeit, Millionen von Variablen gleichzeitig zu analysieren und daraus Muster abzuleiten, übertrifft das menschliche Denkvermögen bei Weitem.

Die Zusammenarbeit zwischen Mensch und KI eröffnet hier jedoch spannende Möglichkeiten. KI kann als Werkzeug dienen, das das menschliche Erfassungsvermögen ergänzt und dem Menschen somit hilft, komplexe Systeme zu verstehen oder zumindest vorherzusagen.

Innovationsplattform KI

Die Verwendung von KI in Simulationen markiert einen Wendepunkt in der Industrialisierung. Simulationen sind längst nicht mehr nur ein Werkzeug, um bestehende Prozesse zu analysieren. Sie haben sich zu einer Innovationsplattform entwickelt, auf der neue Industrien und Technologien entstehen. Insbesondere die Kombination von KI mit weiteren Technologien birgt ein enormes Potenzial, unsere Welt zu transformieren.

Die Industrialisierung der KI bedeutet, dass sie nicht mehr nur experimentell oder punktuell eingesetzt wird. Vielmehr wird sie zu einem integralen Bestandteil von Produktions- und Entwicklungsprozessen. Dies erfordert jedoch einen Wechsel von allgemeinen KI-Architekturen hin zu spezifischen, maßgeschneiderten Modellen. Diese spezialisierten Systeme können für einzelne Branchen oder Anwendungen optimiert werden, sei es für die Optimierung von Energieflüssen, die Fertigung in der Automobilindustrie oder die Steuerung komplexer logistischer Netzwerke.

Ein weiteres Beispiel ist die Automobilindustrie. Hier werden Simulationen eingesetzt, um Fahrzeugsicherheit, Kraftstoffeffizienz und aerodynamische Eigenschaften zu optimieren. Mit KI lassen sich virtuelle Crashtests durchführen, indem man die Materialverformung simuliert, was kostengünstiger als physische Tests ist. Dies beschleunigt die Entwicklung neuer Fahrzeugdesigns und -materialien und macht Fahrzeuge gleichzeitig sicherer.

Die Zukunft der Simulation liegt in der Entwicklung maßgeschneiderter KI-Modelle, die für spezifische Anwendungen optimiert werden. In der ersten Phase der KI-Revolution dominierten universelle Architekturen wie

Transformer-Modelle, die für eine Vielzahl von Aufgaben eingesetzt werden konnten. In der Phase der Industrialisierung wird es jedoch notwendig, diese Architekturen anzupassen, um die spezifischen Anforderungen einzelner Branchen zu erfüllen.

Eine der größten Herausforderungen bei jeder Simulation ist es, die richtige Balance zwischen Schnelligkeit und Genauigkeit zu finden. Die KI muss lernen, wie man komplexe Systeme mit weniger Komponenten beschreibt, ohne die Genauigkeit zu verlieren. Oft wird dabei mit Hidden Spaces gearbeitet. Das bedeutet, dass die KI die Welt auf eine kleinere Anzahl von Parametern reduziert, die die wesentlichen Merkmale und Muster des Systems beschreiben. In diesen abstrakten Räumen kann die KI dann die physikalischen Prozesse simulieren, was viel schneller geht, als wenn sie jedes einzelne Atom berücksichtigen müsste.

Der Hidden Space ist ein abstrakter Raum, in den die Realität projiziert wird. Anstatt Millionen von Atomen eines Kunstdüngerkorns zu simulieren, reduziert die KI das System auf eine vereinfachte Darstellung. Diese Abstraktion ermöglicht es, das Verhalten der Körner schneller zu simulieren. Die KI lernt, wie diese Körner im nächsten Zeitschritt interagieren, und kann die Simulation dann auf diese höhere Ebene beschränken. Dadurch sparen wir extrem viel Rechenleistung und Zeit.

Ein faszinierender Aspekt der KI-basierten Simulation ist, dass die KI Makrostrukturen identifizieren kann, ohne die zugrunde liegenden physikalischen Gesetze im Detail zu verstehen. Die KI extrahiert Muster und Regelmäßigkeiten aus den Daten und nutzt diese, um Vorhersagen zu treffen. Das Ergebnis ist eine Art »neue Physik«, die nicht auf theoretischen Gleichungen basiert, sondern auf empirischen Beobachtungen.

Zum Beispiel könnte die KI lernen, wie sich Sandkörner in einer Maschine verhalten, indem sie riesige Datenmengen analysiert. Sie erkennt Muster, etwa wie Körner sich bei unterschiedlichen Feuchtigkeitsgraden oder unter Druck bewegen. Diese Erkenntnisse kann die KI dann auf neue Geometrien der Maschine anwenden, ohne dass eine vollständige physikalische Modellierung nötig wäre. Oder es wird die Größe der

Sandkörner oder deren Feuchtigkeit verändert. Die KI kann auch dann noch simulieren, wenn sie an vielen Daten mit verschiedener Körnergröße oder Feuchtigkeit trainiert wurde. Das ist eben die Generalisierungsfähigkeit der KI: Sie kann das Gelernte auf unbekannte Parameter übertragen, ohne neu lernen zu müssen.

Die KI abstrahiert das Getreidekorn zu einem Objekt mit bestimmten Parametern wie Größe, Gewicht und Form. Darüber hinaus kann die KI sogar lernen, größere Strukturen wie Kornansammlungen oder Strömungen als Makroobjekte zu betrachten. Diese Art der Abstraktion ermöglicht es, die Simulation nicht nur schneller, sondern auch skalierbarer zu machen. Die KI-Simulation lernt die Wechselwirkungen dieser größeren Strukturen. Dadurch können wir Simulationen viel schneller und effizienter durchführen, ohne viel an Genauigkeit zu verlieren.

Ein weiterer entscheidender Vorteil der KI ist ihre Fähigkeit, aufgrund ihrer Schnelligkeit in Echtzeit zu arbeiten, wo numerische Simulationen versagen. Während numerische Simulationen oft Monate dauern, kann eine KI-gestützte Simulation bis zu tausendmal schneller sein. Dies macht es möglich, bestimmte Simulationen parallel zu einem laufenden Prozess auszuführen – ein Konzept, das als »digitaler Zwilling« bekannt ist.

7

Digitale Zwillinge

KI-gestützte Simulationen können Prozesse nicht nur offline simulieren, um sie zu verbessern oder anzupassen, sondern sie auch in Echtzeit simulieren, um sie zu steuern. Ein Beispiel sind digitale Zwillinge – virtuelle Abbilder realer Systeme, die in Echtzeit aktualisiert werden. Ein solcher digitaler Zwilling könnte etwa den chemischen Prozess einer Produktionsanlage nachbilden, wodurch man eingreifen kann, bevor etwas Unerwünschtes geschieht. Unter digitalen Zwillingen verstehe ich hier also Echtzeitsimulationen von Prozessen und Abläufen. Wenn ein Prozess läuft, dann wird der gleiche Prozess am Computer imitiert, aber eventuell ein wenig schneller. Die Echtzeitsimulation wird dabei kontinuierlich mit Messdaten aus der Realität abgeglichen. Dadurch bleibt der digitale Zwilling nahe am realen Prozess und weicht nicht über die Zeit hinweg ab. Es kann dann nicht nur der aktuelle Zustand des Systems dargestellt werden, sondern auch zukünftige Entwicklungen können vorhergesagt werden.

Digitale Zwillinge könnten in nahezu allen Bereichen unseres Lebens Anwendung finden: Produktionsprozesse und chemische Prozesse könnten gesteuert werden, was in der Medikamentenherstellung oder der Lebensmittelproduktion relevant ist, aber die Möglichkeiten sind nahezu unbegrenzt. Mit der kontinuierlichen Weiterentwicklung von KI und Sensortechnologie wird der digitale Zwilling zu einem unverzichtbaren Werkzeug in Wissenschaft, Industrie und Gesellschaft.

Digitale Zwillinge bieten immense Chancen, Prozesse zu optimieren, Innovationen voranzutreiben und gleichzeitig Kosten zu senken – in einer Welt, die zunehmend auf Daten und Effizienz setzt, sind sie ein unverzichtbarer Bestandteil der Zukunft. Die Geschwindigkeit, mit der KI

arbeitet, ist dabei ein entscheidender Faktor für digitale Zwillinge. Während klassische numerische Simulationen oft zu lange brauchen, um komplexe Prozesse in Echtzeit abzubilden, können KI-Modelle in bestimmten Fällen Prozesse in Echtzeit simulieren. Diese Echtzeitfähigkeit der KI eröffnet neue Perspektiven, insbesondere in Branchen, in denen es ein großer Vorteil ist, einen Prozess zu begleiten.

Beispielsweise könnte ein digitaler Zwilling einer Produktionsanlage in Echtzeit erkennen, wenn eine Maschine ineffizient arbeitet oder die Qualität des Produktes geringer wird. Wenn etwa eine chemische Reaktion außer Kontrolle gerät, könnte nicht nur die Produktionsanlage beschädigt werden, sondern auch die Qualität des Endprodukts leiden. Ein digitaler Zwilling, der in Echtzeit arbeitet und den Produktionsprozess um Sekunden vorausberechnet, kann solche Szenarien verhindern. Dieser virtuelle Zwilling simuliert den Prozess basierend auf aktuellen Daten und ermöglicht Eingriffe, bevor eine Eskalation stattfindet.

Ein Hochofen ist hierfür ein gutes Beispiel: Um zu verstehen, wie sich die Eisenschmelze verhält, müssen wir wissen, wie sich die Partikel im Ofen verhalten, wie die Temperaturen steigen und wie die chemischen Reaktionen ablaufen. Solche Simulationen erfordern komplexe Berechnungen, die viel Zeit und Ressourcen benötigen. Eine KI könnte hier helfen, indem sie Muster in den Daten erkennt und schneller zu Ergebnissen kommt. Ein digitaler Zwilling, der den Betrieb eines Hochofens in Echtzeit begleitet, könnte potenzielle Probleme wie Überhitzung oder Materialstau frühzeitig erkennen und Gegenmaßnahmen vorschlagen.

Digitale Zwillinge können Fehler frühzeitig erkennen und beheben, bevor sie in der Realität auftreten. In der Energieerzeugung könnten Simulationen helfen, sofortige Anpassungen vorzunehmen, um Schwankungen in der Nachfrage oder Produktion auszugleichen. Besonders bei erneuerbaren Energien ist diese Flexibilität entscheidend, wenn das Wetter und die Turbinen simuliert werden und somit frühzeitig auf Wetterschwankungen reagiert werden kann. So könnte ein digitaler Zwilling einer Windturbine genutzt werden, um in Echtzeit zu analysieren, wie man

die Effizienz einer Turbine steigert, indem man sie an die Windverhältnisse anpasst. Digitale Zwillinge können auch Teil von Modellen ganzer Städte sein, um zum Beispiel die Energieverteilung effizienter zu gestalten, wenn Transformatoren simuliert werden, oder um Umweltbelastungen zu minimieren, wenn die Luft- und Abgasströme simuliert werden.

Eines der beeindruckendsten Potenziale von KI-gestützten Simulationen zeigt sich in hochsensiblen Bereichen wie der Kernfusion. Bei diesen Prozessen können kleinste Fehler katastrophale Auswirkungen haben. Hier ist es beispielsweise essenziell, dass das Plasma durch magnetische Felder stabil an einem Ort gehalten wird. Sollte es die Randbereiche des Reaktors berühren, könnten Schäden an der Anlage entstehen. Dabei spielt KI eine entscheidende Rolle: Simulationen können den Prozess in Echtzeit begleiten und kritische Zustände vorhersagen, um präventive Maßnahmen zu ergreifen, bevor Schäden entstehen.

Eine der größten Herausforderungen bei digitalen Zwillingen besteht darin, geeignete Trainingsdaten zu generieren. Reale Daten aus industriellen Prozessen sind oft schwer zugänglich, da Messungen in den innersten Bereichen eines Hochofens oder Reaktors technisch nicht möglich sind. Daher greift man auf Daten aus numerischen Simulationen zurück. Diese Daten werden mithilfe von physikalischen Gesetzen oder fundierten Annahmen erzeugt und liefern eine Datenbasis, mit der eine KI trainiert werden kann. Die Datenerzeugung ist natürlich langwierig, aber sie lohnt sich.

Echtzeitsimulation

Die Entwicklung und Implementierung digitaler Zwillinge ist jedoch nicht ohne Herausforderungen. Die Verarbeitung riesiger Datenmengen in Echtzeit erfordert Rechenressourcen. Zudem müssen Daten sicher und datenschutzkonform verarbeitet werden, um Missbrauch zu verhindern. Eine weitere Herausforderung liegt in der Standardisierung: Unterschiedliche

Systeme müssen miteinander kompatibel sein, damit die Daten zu einem Trainingsdatensatz zusammengefasst werden können.

Echtzeitsimulationen benötigen kontinuierlich aktuelle Daten. Die Erhebung und Verarbeitung dieser Daten stellt eine logistische und technische Herausforderung dar. Die Verarbeitung der enormen Datenmengen erfordert leistungsfähige Hardware und spezialisierte Algorithmen. Insbesondere muss die Hardware robust gegen Einflüsse der Umgebung, etwa Luftfeuchtigkeit, Staub, Vibrationen und Ähnliches, sein. Zudem müssen KI-Modelle eine Balance zwischen Abstraktion und Genauigkeit finden. Zu detaillierte Modelle sind ineffizient, zu abstrakte Modelle können wichtige Details übersehen.

Hinzu kommen bei Echtzeitsimulationen hohe Anforderungen an effiziente Modellarchitekturen. Eine typische physikalische Simulation, etwa zur Analyse von Luftströmungen über einer Flugzeugtragfläche, erfordert die Berechnung von Millionen von Gitterpunkten. Jeder dieser Punkte repräsentiert einen Zustand, wie Druck, Temperatur oder Geschwindigkeit, der kontinuierlich aktualisiert werden muss. Dies macht klassische numerische Methoden langsam und für Echtzeitanwendungen oft unpraktisch. Insbesondere wenn die Gitterpunkte nicht regelmäßig verteilt sind, werden numerische Methoden langsam, da sie die Fast-Fourier-Transformation nicht anwenden können. KI-Methoden können jedoch mit unregelmäßig verteilten Gitterpunkten umgehen. Statt Gitterpunkten können auch Einzelpartikel und deren Bewegung mit KI schneller simuliert werden.

Ebenso können KI-Modelle zeitliche Sprünge machen, indem sie lernen, welche Zustände vorhersehbar sind und welche detailliertere Berechnungen erfordern. So lassen sich Simulationen schneller und dennoch präzise durchführen, selbst bei hochkomplexen Systemen.

Digitale Zwillinge versus Weltmodelle

In diesem Zusammenhang ist es notwendig, noch einmal auf das bereits weiter oben skizzierte Konzept der Weltmodelle zurückzukommen. Im Gegensatz zur Simulation ist die Modellierung allgemeiner, da sie sich nicht nur auf physikalische oder chemische Prozesse beschränkt. Simulation kann als eine physikalische Modellierung gesehen werden. Verglichen mit Simulationen sind Weltmodelle meist nicht sehr exakt und dienen vor allem dazu, zielgerichtete Aktionen auszuführen. Diese Aktionen sollen die Welt in gewünschter Weise beeinflussen oder helfen, bestimmte Ziele zu erreichen. Die Simulation fokussiert sich auf zeitliche physikalische Prozesse der realen Welt, die exakt und schnell approximiert werden sollen, um die Prozesse zu optimieren oder zu steuern. Hingegen sind Weltmodelle immer mit einem Agenten verbunden, der sie benutzt, um ein bestimmtes Ziel zu erreichen. Weltmodelle sagen vorher, was unsere Aktionen in der Welt auslösen, ob unsere Aktionen zielführend sind, ob sie uns schaden oder helfen und wie viel Gewinn oder Verlust wir am Ende machen. Wie schon gesagt, haben auch Tiere ein Weltmodell, das je nach Intelligenz komplexer oder einfacher ist.

Die meisten Weltmodelle bilden eine Repräsentation der Welt und sagen dann vorher, was in der realen Welt geschehen wird. Auch Large Language Models (LLMs) können als textbasierte Weltmodelle gesehen werden. Solche LLMs, wie GPT, abstrahieren die physikalische Welt und bringen sie in eine Form, die für Menschen verständlich ist, da sie ja Texte generieren. Ein textbasiertes Modell erklärt, warum ein Glas fällt, zerbricht und welche Umstände dies begünstigen.

Ein gutes Weltmodell inkludiert auch ein Modell des Agenten selbst, der es benutzt, denn dieser ist auch Teil der Welt. Meiner Ansicht nach ist Bewusstsein die Folge eines guten Weltmodells, das auch ein Modell von sich selbst enthält. Je besser das Modell von einem selbst, desto klarer und deutlicher ist man sich seiner selbst bewusst. Dann weiß man, wie man die Welt beeinflusst und wie die Welt einen selbst beeinflusst. Man weiß, was man kann und was man nicht kann, was einem guttut und was nicht. Konsequenterweise haben Tiere, je nachdem, wie gut ihr Weltmodell und

ihr Modell von sich selbst ist, verschiedene Grade an Bewusstsein. Intelligentere Tiere sind sich ihrer selbst mehr bewusst, da sie ein besseres Weltmodell und ein besseres Modell von sich selbst haben. Also ist Bewusstsein ein Nebeneffekt eines guten Weltmodells.

Manche Weltmodelle können verschiedene Simulationen kombinieren. Ein solches Modell könnte nicht nur präzise Vorhersagen aufgrund von Simulationen treffen, sondern auch erklären, warum etwas passiert. Ein Beispiel: Wenn wir wissen wollen, wie ein bestimmtes Material auf Hitze reagiert, könnte eine Simulation nicht nur die physikalischen Eigenschaften des Materials simulieren, ein Textmodell könnte auch sprachlich beschreiben, welche Konsequenzen dies für die Praxis hat – etwa für die Produktion von Batterien oder Solarpanels.

Oder ein Weltmodell könnte nicht nur die molekularen Wechselwirkungen eines neuen Medikaments simulieren, sondern auch beschreiben, wie diese Wechselwirkungen die Wirksamkeit und eventuelle Nebenwirkungen beeinflussen. Ein solcher Ansatz würde die Wissenschaft auf eine neue Ebene heben.

8

Komplexität beherrschen

KI wird eine entscheidende Rolle bei der Lösung komplexer Probleme spielen, die wir heute noch nicht vollständig verstehen. Die Fähigkeit, die Welt auf abstrakte und vereinfachte Weise zu modellieren, wird uns in vielen Bereichen voranbringen.

Simulationen könnten zum Beispiel dazu beitragen, erschwingliche Lebensmittel für eine wachsende Weltbevölkerung bereitzustellen. Ebenso könnten Lagerungsprozesse so optimiert werden, dass keine Produkte verderben oder Ressourcen verschwendet werden.

Die Zukunft der Modellierung mithilfe von Simulationen liegt in ihrer umfassenden Integration in alle Bereiche der Industrie und Gesellschaft, von Produktion und Maschinenbau über Medizin bis zur globalen Klimamodellierung. Komplexe Systeme wie Energiesysteme oder das Klima stellen die Menschen dabei vor gewaltige Herausforderungen, die ohne KI-Weltmodelle wohl nicht bewältigt werden können.

Solche Weltmodelle können aus vielen interagierenden Teilsystemen bestehen, deren Verhalten oft nicht linear ist. Kleine Veränderungen können massive Auswirkungen haben – oder auch keine, je nach der Art der Interaktion. Ein einfaches Beispiel ist das Wetter. Das Wetter hängt auch von Faktoren wie Bodenzustand, Vegetation oder unterirdischen Wasserströmen ab, wofür man dann Modellierungen verwenden muss, da es hierfür keine Physik gibt, die man simulieren kann. Ähnliche Dynamiken, die durch Modellierung beschrieben werden können, finden sich in Energiesystemen, wo Verbrauchsmuster von Menschen, die Verfügbarkeit erneuerbarer Energien und technische Störungen miteinander verknüpft sind.

Um solche Systeme zu modellieren, reicht es nicht, ein KI-Modell für das Gesamtsystem zu bauen. Stattdessen müssen die Teilsysteme separat verstanden und optimiert werden. Für mich ist das ein zentrales Merkmal von komplexen Systemen: die Interaktion der Komponenten und ihre Nichtlinearität. Besonders spannend wird es, wenn sich kleine Änderungen in der Eingabe zu riesigen Unterschieden im Output entwickeln – etwa in der Medikamentenentwicklung, wenn eine minimale chemische Modifikation ein Molekül plötzlich hochwirksam macht.

Ich glaube fest daran, dass KI-Modellierung vor allem dafür eingesetzt werden sollte, uns bei der Bewältigung der Probleme mit komplexen Systemen zu helfen. Ob es um den Klimawandel, die globale Energieversorgung oder die Logistik in einer immer mehr vernetzten Welt geht – die Herausforderungen, vor denen wir stehen, übersteigen oft unsere menschlichen Fähigkeiten. Hier kann KI uns unterstützen, indem sie Muster erkennt, die wir nicht sehen, und Lösungen vorschlägt, die wir allein nicht entwickeln könnten.

Trotz aller Fortschritte gibt es allerdings auch Grenzen. Modelle und Simulationen sind immer nur so gut wie die Daten, auf denen sie basieren. In vielen Bereichen fehlen uns jedoch noch die notwendigen Daten, um wirklich präzise Modelle zu erstellen. Beispielsweise stoßen wir bei biologischen Systemen an solche Grenzen, etwa bei der Simulation von Blutflüssen in Adern oder Herzen. Dies liegt daran, dass biologische Systeme wie die Zelle, das Gehirn, das Immunsystem oder verschiedene Organe so komplex sind, dass wir sie nicht gut modellieren können.

Ein weiteres Hindernis ist die Rechenleistung. Obwohl moderne Supercomputer beeindruckende Fähigkeiten haben, reichen sie oft nicht aus, um extrem große Simulationsmodelle durchzurechnen. Die Interpretation von Simulations- und Modellierungsergebnissen erfordert immer noch menschliches Urteilsvermögen, da KI nicht in der Lage ist, Kontext oder ethische Fragestellungen zu berücksichtigen.

Intelligente Landwirtschaft

Die Wirtschaft ist ein extrem komplexes System, das von unzähligen Faktoren beeinflusst wird: von menschlichem Verhalten über geopolitische Ereignisse bis hin zu Wetterbedingungen. Simulationen funktionieren gut in physikalischen Bereichen wie der numerischen Strömungssimulation oder der Optimierung von Produktionsprozessen. Doch sobald menschliches Verhalten ins Spiel kommt, wird es schwierig, und man kann nur mehr modellieren. Die Modellierung ist hier allerdings ebenfalls sehr schwierig, da meist Daten fehlen und komplexe, vielfältige, nichtlineare Interaktionen stattfinden.

Ein Hindernis im Bereich der Wirtschaft ist der Mangel an Daten. Während wir in der Industrie oft umfangreiche historische oder simulierte Daten haben, die als Grundlage für das Lernen eines Simulationsmodells dienen können, fehlen solche Daten zum Lernen von Modellen in der Wirtschaft. Jedes Jahr bringt neue Variablen und Rahmenbedingungen mit sich – sei es durch technologische Fortschritte, geopolitische Veränderungen oder gesellschaftliche Trends. Das macht es nahezu unmöglich, Modelle zu entwickeln, die zuverlässig genug sind, um komplexe wirtschaftliche Entscheidungen zu modellieren. Das Problem ist, dass die Wirtschaft nicht stationär ist, das heißt, die Dynamiken verändern sich laufend, sodass historische Daten nicht für die Zukunft helfen.

Ein weniger beachtetes Feld, in dem KI große Veränderungen bewirken könnte, ist die Landwirtschaft, die auch sehr komplexe Systeme beinhaltet. Hier werden bereits heute autonome Mähdrescher und Erntemaschinen eingesetzt, um große Gebiete abzuernten. In den USA gibt es sogar Mähdrescher ohne Führerhaus, die ferngesteuert über riesige Felder fahren, oder Drohnen, die Äpfel und Birnen pflücken. Solche Technologien machen die Arbeit effizienter und präziser.

Mit Drohnen können auch Felder überwacht werden, um genau festzustellen, wo Dünger oder Schädlingsbekämpfungsmittel benötigt

werden. Statt pauschal das gesamte Feld zu behandeln, wird gezielt dort eingegriffen, wo es nötig ist. Das spart Ressourcen und schont die Umwelt. Unternehmen wie Bayer arbeiten bereits an solchen Technologien, die ein enormes Potenzial haben.

Auch im Weinbau könnte man KI und Modellierung nutzen, um basierend auf Wetterdaten, Bodenzustand und Reifegrad der Trauben den perfekten Erntezeitpunkt zu berechnen oder die Qualität des Weins vorherzusagen. Doch dafür müsste man über Jahre hinweg detaillierte Daten sammeln – etwa Bilder der Reben, klimatische Bedingungen und chemische Analysen des Bodens. Diese Daten fehlen bisher häufig, insbesondere bei kleineren Betrieben. Große Weingüter könnten solche Systeme jedoch entwickeln, da sie eher in der Lage sind, Sensoren und Technologien zu integrieren.

Ein Aspekt, der hierbei allerdings häufig übersehen wird, ist die kulturelle Bedeutung der Landwirtschaft. Viele Menschen verbinden mit der bäuerlichen Arbeit ein Stück Tradition und Identität. Die Vorstellung, dass der Bauer der Zukunft nur noch am Computer sitzt, ist für manche schwer zu akzeptieren. Dennoch bleibt der Kern der Arbeit bestehen: Entscheidungen müssen weiterhin von Menschen getroffen werden, selbst wenn KI die Analyse und Datenerhebung übernimmt. Ob ein Bauer über ein Feld läuft oder mit einer Drohne Bilder betrachtet, ändert nichts an seiner Expertise.

KI für die Industrie

In der aktuellen Phase der KI-Entwicklung sehen wir eine Verschiebung von der Skalierung hin zur Industrialisierung. Große Modelle wie GPT haben gezeigt, wie leistungsfähig KI sein kann. Doch die nächste Herausforderung besteht darin, diese Modelle für spezifische Anwendungsbereiche anzupassen und weiterzuentwickeln.

Dabei kann es sinnvoll sein, kleinere, robustere Modelle zu entwickeln, die speziell für eine Aufgabe optimiert sind. Beispielsweise könnte ein Modell, das in der Chemieindustrie eingesetzt wird, spezifische Eigenschaften chemischer Reaktionen berücksichtigen, während ein Modell für die Lebensmittelindustrie mit Daten über Lagerungs- und Logistikprozesse arbeitet.

Ein spannender Ansatz sind die bereits erwähnten Foundation Models, die übergreifendes Wissen in verschiedenen Domänen lernen. Sie lernen nicht, eine bestimmte Aufgabe zu lösen, sondern eine Menge von Aufgaben. Stellen Sie sich ein Sand-Foundation-Model vor, das Sandkörner und deren Interaktion simulieren kann – unabhängig von ihren Eigenschaften wie Größe, Form, Oberflächenbeschaffenheit oder Feuchtigkeit. Dieses Model kann auch dann eingesetzt werden, wenn Sandkörner mit neuen Eigenschaften zu simulieren sind. Durch die Generalisierung kann das Simulationsnetz nun auch solche Sandkörner simulieren. Foundation Models sind eben Meister in der Generalisierung, da sie die Dynamik oder Physik der Domäne gelernt haben und nicht spezielle Aufgaben. Die Herausforderung besteht darin, die richtigen Daten zu sammeln und die Modelle so zu trainieren, dass sie robust und vielseitig einsetzbar sind. So müssen in den Trainingsdaten viele verschiedene Eigenschaften vorhanden sein, damit das trainierte Foundation Model auf neue Eigenschaften generalisieren kann. Foundation Models kann es für Simulation oder Modellierung geben – Large Language Models (LLMs) etwa sind Foundation Models für Text, da sie allen verfügbaren Text im Training verwenden. Später werden sie dann auf spezielle Aufgaben angepasst, zum Beispiel mit Texten aus dem Bankensektor.

Die Erfahrungen, die wir in der Hochskalierungsphase gemacht haben, sind entscheidend. Durch die Entwicklung großer Foundation Models wie Transformer für Textgenerierung haben wir viel darüber gelernt, wie Foundation Models funktionieren und welche Daten sie benötigen. Dieses Wissen kann jetzt genutzt werden, um neue Foundation Models effizient zu trainieren, ohne jedes Mal bei null anzufangen.

Die Industrialisierung der KI wird auch neue Geschäftsmodelle und Arbeitsweisen hervorbringen. Dabei geht es nicht nur um die Entwicklung von Modellen, sondern auch um die Schaffung einer Infrastruktur, die es ermöglicht, diese Modelle effizient einzusetzen.

Die Stärke der kommenden KI-Systeme liegt in ihrer Anpassungsfähigkeit. Foundation Models werden als Grundlage dienen, die dann für spezifische Anwendungen angepasst werden können. Diese Anpassungen können neue Datenquellen, zusätzliche Sensoren oder spezielle Anforderungen einer Branche berücksichtigen. Wichtig ist dabei die Zusammenarbeit zwischen verschiedenen Disziplinen. KI-Entwickler, Ingenieure und Branchenexperten müssen kooperieren, um die besten Lösungen zu finden. Das erfordert nicht nur technisches Know-how, sondern auch ein Verständnis für die spezifischen Anforderungen und Herausforderungen der jeweiligen Domäne. Insbesondere beim Trainieren von Foundation Models muss man wissen, welche Eigenschaften, Nebenbedingungen oder geometrischen Beschränkungen in der Praxis Sinn ergeben und als Trainingsdaten dienen sollen. Hier braucht man Domänenexperten, um Trainingsdaten für Foundation Models zu erzeugen.

Schlauere Autos

Die deutsche Autoindustrie steht vor einer massiven Transformation. Die Zeiten, in denen Lebensdauer, Geschwindigkeit und Motorleistung die zentralen Verkaufsargumente waren, neigen sich dem Ende zu. Stattdessen wird das Auto der Zukunft eher ein intelligenter Lebensraum sein – eine Art »mobiles Wohnzimmer«. Kunden wünschen sich Komfort, Personalisierung und intelligente Interaktion. Hier bietet KI unzählige Möglichkeiten, von der Gestaltung des Innenraums bis hin zu Sicherheits- und Unterhaltungssystemen.

Ein Beispiel ist die Integration von KI in Produktionsprozesse. Während Roboter traditionell für die Herstellung großer Serien optimiert sind, wird

KI künftig auch bei individualisierten Serien eine Schlüsselrolle spielen. Nehmen wir kleine Chargen von Autos mit speziellen Anpassungen, sei es bei der Innenausstattung oder dem Design. Mit KI könnte man Roboter so trainieren, dass sie durch wenige Beispiele lernen, diese Anpassungen effizient umzusetzen. Dadurch würden auch Kleinserien wirtschaftlich rentabel. Hier könnte die neue xLSTM-Technologie zum Einsatz kommen.

Meiner Ansicht nach sollte sich die Autoindustrie künftig weniger auf die Hardware konzentrieren, sondern mehr auf digitale Services. Ich habe einmal zu Forschern bei Audi gesagt: »Stellt euch vor, ihr baut nicht nur ein Auto, sondern ein riesiges Smartphone, in das man sich hineinsetzen kann.« Updates, neue Funktionen und individuelle Anpassungen – genau wie bei einem Smartphone – könnten das Kundenerlebnis laufend verbessern. Warum sollte ein Auto im Winter nicht mit einem speziellen Feature ausgestattet werden, das per Software-Update hinzugefügt wird?

Tesla hat gezeigt, wie erfolgreich dieses Modell sein kann. Statt nur auf die Leistung der Fahrzeuge zu setzen, hat Tesla den Fokus auf das Benutzererlebnis und smarte Technologien gelegt. Das ist eine Lehre, die die deutsche Autoindustrie verinnerlichen muss, wenn sie im globalen Wettbewerb bestehen will.

Ein anderes Thema, das immer wieder für Diskussionen sorgt, ist das autonome Fahren. Werden wir jemals vollautonome Autos erleben, die ohne menschliches Eingreifen funktionieren? Vielleicht, aber ich bin skeptisch, ob das in absehbarer Zeit möglich ist.

Um vollständig autonom zu fahren, reicht es nicht, die bisherigen Daten zu nutzen oder Verkehrssituationen zu analysieren. Ein Auto müsste die Welt verstehen, ähnlich wie ein Mensch. Das bedeutet, das KI-System müsste unterscheiden können, ob ein Objekt auf der Straße eine Plastiktüte oder ein Stein ist. Wenn das Objekt verweht wird, dann ist es eher eine Plastiktüte. Aber dazu muss man Weltwissen haben, also wissen: Plastiktüten sind leicht, es gibt Wind, und der Wind verweht Plastiktüten. Das KI-System müsste komplexe soziale und physische Interaktionen interpretieren können. Wenn sich die Menschen einer Gruppe die Hand

geben, dann gehen sie auseinander und eventuell sogar auf die Straße. Wenn ein Ball auf die Straße rollt, könnte ein Kind hinterherlaufen. Wenn sich ein Hund von der Leine losreißt, so wird das Frauchen oder Herrchen hinterherlaufen. Solange wir keine künstliche Intelligenz haben, die ein menschliches Weltverständnis besitzt, sind wir weit davon entfernt, vollautonome Fahrzeuge auf die Straße zu bringen. Ähnlich wie bei Large Language Models zeigt sich, dass diese Systeme trotz beeindruckender Leistungen in vielen Bereichen immer noch grundlegende Grenzen haben.

KI in der Pharmaindustrie

Die Einführung von KI in der Pharmaindustrie verlief ähnlich holprig wie in der Automobilbranche. Vor 18 Jahren, als wir mit ersten Projekten in diesem Bereich starteten, war die Skepsis groß. Viele fragten sich: »Wie soll KI chemische Prozesse vorhersagen?« Doch die jüngere Generation, die während ihres Studiums oder in Forschungsprojekten bereits mit KI gearbeitet hat, hat diese Denkweise verändert. Heute sehen wir, wie stark KI die frühen Phasen der Medikamentenentwicklung beschleunigt, ja geradezu revolutioniert. KI-Modelle können etwa Nebenwirkungen vorhersagen, lange bevor erste klinische Tests stattfinden. Früher war dieser Prozess extrem aufwendig und zeitintensiv. Es mussten Tausende von Molekülen für Zehntausende von Nebenwirkungen getestet werden. Nebenwirkungen wurden oft mithilfe von Tierversuchen untersucht, die nicht nur ethische Fragen aufwarfen, sondern auch teuer und ineffizient waren.

Ein Aha-Moment für viele Pharmaforscher war, als sie erkannten, dass KI-Vorhersagen nicht nur schneller, sondern in einigen Fällen auch genauer waren als traditionelle Tests. Während Laborexperimente von Natur aus Schwankungen unterliegen – etwa durch variierende Umgebungsbedingungen oder Messfehler –, liefern KI-Modelle stabilere Ergebnisse. Das ist besonders bei der Vorhersage von Nebenwirkungen oder der Bindung von Molekülen an Rezeptoren entscheidend.

Dieser Wandel hat nicht nur die Art und Weise verändert, wie Medikamente entwickelt werden, sondern auch den gesamten Forschungsprozess beschleunigt. In den frühen Phasen der Entwicklung können KI-Systeme Tausende von möglichen Molekülen bezüglich Zehntausender Nebenwirkungen durchleuchten und diejenigen identifizieren, die mit hoher Wahrscheinlichkeit erfolgreich sein werden. Oft wird ein Ranking gemacht, sodass die erfolgversprechendsten Moleküle zuerst getestet werden. Das spart Zeit und Ressourcen und reduziert die Anzahl der Moleküle, die in Tierversuchen getestet werden müssen.

Ähnliche Entwicklungen sehen wir auch in anderen Bereichen der Medizin. In der Pharmaindustrie etwa spricht man von »virtuellen Bioassays«, bei denen KI-Systeme auf der Basis von Daten die Wirkung und Nebenwirkungen von Molekülen bestimmen. Virtuelle Bioassays sind sehr erfolgreich in der Pharmaindustrie, wo sie schon einige Bioassays wie Molekülbindungsexperimente, Petrischalenexperimente oder Tierexperimente ersetzt haben. Jedoch mussten vorher genügend Daten aus den ursprünglichen Bioassays verfügbar sein, um ein Modell zu trainieren, das dann als virtuelles Assay gesehen werden kann.

Kampf gegen den Klimawandel

Auf der Basis von LSTM konnten wir hydrologische Modelle entwickeln, die für Flutwarnungen weltweit eingesetzt werden, unter anderem von der US-amerikanischen, der kanadischen und der israelischen Regierung. Diese Modelle waren so präzise, dass Google sie kommerzialisiert hat – eine ihrer Apps gibt Nutzern präzise Vorhersagen darüber, ob eine Flut ihr Zuhause gefährden könnte. Diese Technologie führte zu einem Paradigmenwechsel in der Hydrologie. Hier hat die KI die etablierten Modelle verdrängt. Die klassischen Ansätze, die oft auf jahrzehntelanger Forschung und Modellierung bezüglich Bodenbeschaffenheit, Bepflanzung, Gletscherschmelzen, Schneeschmelzen und Ähnlichem beruhten, wurden

von LSTM-Modellen übertroffen. Diese KI-Systeme konnten aus Daten des Flusswasserstandes und einer Beschreibung des Flussbeckens auch Schneeschmelze, Bodenwasseraufnahme und -abgabe und unterirdische Wasserströme automatisch beschreiben, ohne dass hierfür Daten vorhanden waren. Die Vorhersagen sind präziser als mit herkömmlichen Modellen. Die Genauigkeit der LSTM-Modelle löste einen Schock innerhalb der Hydrologen aus, die nicht glauben konnten, dass datengetriebene Methoden so gut sein konnten.

Viele Hydrologen, die ihr Berufsleben auf traditionelle Modelle gestützt hatten, waren zunächst skeptisch. Einige waren so schockiert von den Ergebnissen, dass sie ihre wissenschaftliche Arbeit vorübergehend pausierten, um sich neu zu orientieren. Andere weigerten sich, die Ergebnisse anzuerkennen, bis umfassende Vergleichsstudien zeigten, dass die LSTM-Modelle in nahezu allen Szenarien besser abschnitten. Dies führte zu einem Paradigmenwechsel von klassischen Modellen hin zu KI-Modellen, die auf LSTMs basieren. Dies zeigt, dass KI nicht nur neue Möglichkeiten eröffnet, sondern auch bestehende Denkmuster und Methoden infrage stellt.

KI kann auch helfen, Prozesse effizienter zu gestalten. Zum Beispiel könnten Wärmepumpen in Häusern intelligenter gesteuert werden. Auch die Entwicklung neuer Materialien, etwa für die CO_2-Abscheidung, könnte durch KI-Simulationen beschleunigt werden. Bereits jetzt gibt es KI-gestützte Simulationen, um Katalysatoren und andere Technologien zu entwickeln, die CO_2 effizient aus der Atmosphäre entfernen.

Die Grenzen der KI zeigen sich bei hochgradig heterogenen Systemen, wie etwa Energiemärkten. Diese Systeme sind nicht nur von physischen und technischen Faktoren wie Wetter oder Infrastruktur abhängig, sondern auch von menschlichem Verhalten, politischen Entscheidungen und wirtschaftlichen Schwankungen. Beispielsweise müsste ein Modell, das den Energieverbrauch während einer Skiweltmeisterschaft prognostiziert, auch die Erfolgserwartung und den derzeitigen Erfolg der nationalen Mannschaft

einbeziehen. Diese Vielzahl von Faktoren macht es schwer, präzise Modelle zu erstellen. Die Modelle müssten sehr viel Weltwissen haben, um auch psychologische Faktoren richtig in ein Modell zu integrieren. Solche Herausforderungen zeigen, dass wir noch weit davon entfernt sind, KI in allen Bereichen effektiv einzusetzen. KI kann eben nicht alles. Man muss auch ihre Grenzen kennen.

Einsätze in der Medizin

Ein weiteres Beispiel für den Einsatz von Simulationen ist die Pandemiebekämpfung. Man könnte das Luft-Viren-Gemisch simulieren, um zu sehen, wie viele Viren bei bestimmten Wetterverhältnissen übertragen werden. Aber meist wird Modellierung verwendet. Während der Covid-19-Pandemie wurde die Ausbreitung des Virus modelliert, um den Effekt verschiedener Maßnahmen zu bewerten. Szenarien wie »Was passiert, wenn alle eine Maske tragen?« oder »Wie wirkt sich eine Schulschließung auf die Infektionsrate aus?« wurden modelliert, um Politikern und Gesundheitsexperten bei ihren Entscheidungen zu helfen. Diese Modelle waren jedoch oft stark vereinfacht und mussten laufend mit neuen Daten angepasst werden, um genauere Ergebnisse zu liefern. Die Weiterentwicklung solcher Modelle auch durch Simulationen könnte in zukünftigen Pandemien entscheidend sein, um verschiedene Maßnahmen hinsichtlich ihrer Effektivität besser zu bewerten.

Wie bereits erwähnt, brauchen Modellierungen und Simulationen jedoch als Grundvoraussetzung Daten, die uns in einigen Forschungsbereichen allerdings noch fehlen. Ein Beispiel für dieses aktuelle »Datenproblem« finden wir erneut in der Biologie bzw. in der Medizin: Eine Krebszelle so detailliert zu modellieren, dass wir am Computer potenzielle Medikamente testen und ihre Wirkung vorhersagen können – das ist eine faszinierende Vorstellung. Doch davon sind wir noch sehr weit weg. Das Hauptproblem ist die Komplexität. Eine Zelle ist ein hochgradig vernetztes System mit

Tausenden von Akteuren wie den Organellen: Zellkern, Endoplasmatisches Retikulum, Ribosomen, Golgi-Apparat, Mitochondrien, Vesikel, Plastiden (z. B. Chloroplasten) oder Vakuole. Auf der Zelloberfläche gibt es Kanäle wie Ionenkanäle, Adhäsionskontakte und die extrazelluläre Matrix, wodurch Zellen mit anderen Zellen interagieren. Auf molekularer Ebene interagieren Proteine, DNA-Moleküle, RNA-Moleküle, Peptide, Neurotransmitter, Ionen und viele andere Stoffe. Selbst wenn wir in der Lage wären, die Position und Funktion jedes dieser Elemente zu messen, hätten wir immer noch das Problem, dass wir ihre Interaktionen verstehen müssten. Und das tun wir bislang nicht.

Während wir die grundlegenden physikalischen Prinzipien der Schwerkraft genau verstehen, fehlt uns dieses Wissen für die Prozesse innerhalb einer Zelle. Wir haben keine mathematischen Modelle, die alle relevanten Interaktionen – von Proteinen bis hin zu genetischen Prozessen – beschreiben können. Zudem fehlen uns die Daten, um solche Modelle zu validieren.

Ein weiteres Hindernis ist die Skalierung. Selbst wenn wir die Mechanismen einzelner Moleküle verstehen, scheitert eine Modellierung oft an der schieren Menge der Daten, die verarbeitet werden müssten. Eine vollständige Modellierung einer einzigen Zelle auf atomarer Ebene – einer atomaren Simulation also – würde unvorstellbare Rechenressourcen erfordern. Solche Ressourcen stehen uns heute nicht zur Verfügung.

Ein Feld jedoch, wo KI-Simulationen eine immer größere Rolle spielen, ist die Molekulardynamik, ein Forschungsbereich, in dem Wissenschaftler das Verhalten von Molekülen simulieren. Diese Simulationen sind besonders wichtig für Chemie, Biologie und Materialwissenschaft, weil sie helfen, komplexe Prozesse zu verstehen, die in der Natur auf atomarer Ebene ablaufen. Für solche Simulationen werden sogenannte »Force Fields« (Kraftfelder) verwendet, die bereits Makrostrukturen aus mehreren Atomen enthalten. Man muss also nicht jedes einzelne Atom simulieren, was extrem aufwändig wäre. Heute werden die besten Force Fields mithilfe von KI erstellt. Genau darin liegt die Stärke der KI: Sie findet präzisere und

abstraktere Repräsentationen. Indem sie Atome zu Makrostrukturen zusammenfasst, ermöglicht sie effizientere Simulationen.

Ein Hoffnungsschimmer ist die zunehmende Automatisierung in der Biotechnologie. Projekte wie AlphaFold haben gezeigt, dass KI enorme Fortschritte erzielen kann, wenn genügend hochwertige Daten verfügbar sind. AlphaFold konnte die 3D-Strukturen von Proteinen vorhersagen, weil es auf riesigen Datensätzen aus biologischen Experimenten trainiert wurde. Für die Entwicklung von AlphaFold wurde Demis Hassabis der Chemienobelpreis verliehen. Ähnliche Ansätze könnten in der Zukunft auch für Zellen genutzt werden, wenn es gelingt, die notwendigen Daten in ausreichender Quantität und Qualität zu sammeln.

Revolution in der Wissenschaft

Eine spannende Frage ist, ob KI eines Tages eigenständig Forschung betreiben kann – also nicht nur als Werkzeug, sondern als unabhängige Akteurin im wissenschaftlichen Prozess. In gewissem Maße geschieht das bereits. Es gibt KI-Systeme, die wissenschaftliche Literatur analysieren, Trends identifizieren und sogar Vorschläge für neue Experimente machen. Mario Krenn, einer meiner ehemaligen Kollegen, hat beispielsweise ein System entwickelt, das alle Veröffentlichungen zur Quantenphysik analysiert, um vorherzusagen, in welche Richtung sich das Feld entwickeln wird. Solche Ansätze zeigen, dass KI in der Lage ist, nicht nur bestehende Probleme zu lösen, sondern auch neue Forschungsrichtungen zu erschließen.

Doch das allein reicht nicht aus. Für echte autonome Forschung müsste eine KI in der Lage sein, Hypothesen experimentell zu testen und zu validieren. Das kann die KI aber noch nicht. Auch die Relevanz der Ergebnisse in einem gesellschaftlichen Kontext bewerten kann eine KI nicht. Hier liegt die größte Herausforderung. Forschung erfordert auch ein tiefes Verständnis für gesellschaftliche Bedürfnisse, ethische

Fragestellungen und die Bedeutung von Fortschritt. Sonst kann Forschung leicht irrelevante Ergebnisse erzielen, wobei auch wir Menschen die Relevanz schlecht einschätzen können. Schon der berühmte Mathematiker Carl Friedrich Gauß glaubte, dass er seine Entdeckung der komplexen Zahlen nicht veröffentlichen sollte, da diese zu weit ginge. Später wurden die komplexen Zahlen ein sehr wichtiges Instrument in vielen Forschungsfeldern.

Die Zukunft liegt in der Zusammenarbeit zwischen Mensch und Maschine. KI kann komplexe Berechnungen und Analysen durchführen, während Menschen Empathie, Ethik, Risikoabschätzungen und das gesellschaftliche Urteilsvermögen beisteuern, um diese Ergebnisse in sinnvolle Handlungen umzusetzen. Das menschliche Element bleibt aus meiner Sicht unersetzlich.

Ein Beispiel aus der Medikamentenentwicklung verdeutlicht dies: Ein KI-System könnte theoretisch so programmiert werden, dass es automatisch neue Medikamente entwickelt. Doch wer entscheidet, welche Krankheiten priorisiert werden? Wer bestimmt, ob die Forschung auf seltene genetische Erkrankungen oder weit verbreitete Krankheiten wie Diabetes ausgerichtet ist? Solche Entscheidungen erfordern menschliches gesellschaftliches Urteilsvermögen, ethische Einsichten, Empathie und oft auch politisches Geschick.

Eine realistischere Vision für die Zukunft ist aus meiner Sicht, dass KI die Rolle eines Mitarbeiters übernimmt. Sie kann Forscherinnen und Forscher entlasten, indem sie Routineaufgaben übernimmt, Simulationen schneller durchführt oder große Datenmengen analysiert. Dadurch bleibt den Menschen mehr Zeit für die wirklich kreativen und richtungsweisenden Aspekte der Forschung. Ein gutes Beispiel hierfür ist die Astronomie, wo KI bereits genutzt wird, um riesige Mengen an Teleskopdaten zu durchforsten und neue Sterne oder Planeten zu entdecken. Die endgültige Interpretation dieser Entdeckungen bleibt jedoch beim Menschen.

Ähnlich verhält es sich in der industriellen KI. Man könnte ein System entwickeln, das mithilfe von Simulationen verschiedene

Produktionsprozesse optimiert. Aber die Frage, ob diese Simulationen tatsächlich für die Firma Vorteile bringen, kann nur ein Mensch beantworten. Das »Warum« hinter der Forschung bleibt also eine zutiefst menschliche Frage.

9

Chance für Europa

Europa liegt im globalen KI-Wettlauf zurück. Das kann sich aber rasch ändern, wenn wir jetzt das Richtige tun. Das Drei-Phasen-Modell der technologischen Entwicklung (Grundlagenforschung, Skalierung, Industrialisierung – siehe Kapitel 5) kann uns helfen zu verstehen, wo und wie wir in Europa die riesigen Chancen nutzen können, die sich uns bieten.

Die erste Phase war die Grundlagenforschung. In dieser Zeit haben wir die Basis für alles gelegt, was heute als KI gilt. Europa hat sich in dieser Phase sehr gut geschlagen. Viele Technologien, die heute in KI-Systemen zum Einsatz kommen, wurden hier entwickelt. Symbolische KI, Expertensysteme, Logiksysteme – das war unser Ding. Auch das von mir entwickelte LSTM, eines der Herzstücke von Deep Learning, ist eine europäische Erfindung. Später kamen dann weitere Entwicklungen wie die Support-Vector-Maschinen dazu. Diese Grundlagen waren essenziell und haben Europa damals gut positioniert.

Dann kam die zweite Phase: die Hochskalierung. Und hier, das muss ich ehrlich sagen, hat Europa massiv den Anschluss verpasst. Während in den USA große Firmen wie Google, Microsoft oder Meta und in China Baidu, Tencent oder Alibaba die Entwicklung riesiger KI-Modelle vorangetrieben haben, stand Europa abseits. Warum? Ganz einfach: Es fehlte an großen Technologieunternehmen, die in der Lage gewesen wären, solche Modelle zu bauen und zu finanzieren. Auch fehlte es in Europa an den entsprechenden Daten, die die amerikanischen IT-Giganten auf ihren Plattformen automatisch einsammelten. Wenn man selbst nicht »Hand anlegen« kann, sieht man auch nicht, was schiefgeht – und kann daraus

keine neuen Erfindungen machen. Ohne diese großen Modelle fehlt uns in Europa heute eine entscheidende Grundlage, um konkurrenzfähig zu sein.

Das spiegelt sich auch in der Industrialisierung der KI, der dritten Phase, wider, in der wir uns gerade befinden. Es geht jetzt darum, KI in die Breite zu bringen – in Unternehmen, in Produktionsprozesse, in alltägliche Anwendungen. Und wieder: Die USA und China sind uns hier weit voraus. Sie haben KI nicht nur entwickelt, sondern auch in ihre Unternehmen integriert. Denken wir an selbstfahrende Autos. Tesla, Waymo oder Apple haben gezeigt, wie KI in der Automobilbranche eingesetzt werden kann. Die europäischen Hersteller – Mercedes, BMW, Audi – wurden regelrecht gezwungen, sich anzupassen. Aber der Impuls kam von außen, nicht von innen. Europa hat hier nicht nur den Anschluss verpasst, sondern kämpft auch mit strukturellen Hindernissen, die eine schnelle Adaption von KI erschweren.

Europa steht im Bereich der KI vor massiven Herausforderungen, insbesondere bei der Kommerzialisierung und den großen Modellen. Oft heißt es, Europa liege etwa fünf Jahre hinter den USA zurück. Wenn wir diesen Rückstand nicht weiter vergrößern wollen, sind tiefgreifende Veränderungen nötig – sowohl in der Forschungslandschaft als auch in der Zusammenarbeit zwischen Universitäten und Unternehmen.

Eine der Hauptursachen für diesen Rückstand liegt in der fehlenden Durchlässigkeit zwischen akademischer Forschung und der Industrie. Anders als in den USA, wo Forscher oft problemlos zwischen Universitäten und Unternehmen wechseln können, gibt es in Europa immer noch strikte Trennlinien.

Meine Doktoranden stehen oft vor der Entscheidung: »Bleibe ich in der Wissenschaft oder wechsle ich in die Industrie?« Diese Trennung verhindert einen kontinuierlichen Austausch von Wissen und Ideen. In den USA ist die Situation deutlich flexibler. Dort können Wissenschaftler problemlos zwischen Uni und Unternehmen wechseln. Professoren gründen Firmen, Studenten arbeiten an akademischen Projekten und Start-ups – und das alles in einem dynamischen, durchlässigen System.

Durchlässige Strukturen

Was mir immer wieder auffällt, ist die völlig unterschiedliche Einstellung zu Innovationen und Unternehmertum. In den USA ist es selbstverständlich, dass ein Forscher oder Ingenieur nach einer bahnbrechenden Erfindung darüber nachdenkt, ein Start-up zu gründen. Die Aussicht, aus der eigenen Forschung eine Firma aufzubauen, wird dort hochgeschätzt – sowohl gesellschaftlich als auch wirtschaftlich.

Hier in Europa sieht das ganz anders aus. Ein Wissenschaftler, der eine brillante Erfindung macht, denkt oft eher daran, sie zu publizieren und sich einen Namen in der akademischen Welt zu machen. Ein Start-up zu gründen, wird hier nicht mit demselben Respekt betrachtet wie eine Karriere als Professor. Das zeigt sich nicht nur an den persönlichen Ambitionen, sondern auch daran, wie unsere Gesellschaft unternehmerische Risiken wahrnimmt. Ein Gründer in Europa hat es oft schwerer, Unterstützung und Anerkennung zu finden, während in den USA das Scheitern eines Start-ups fast als Zeichen für Erfahrung gilt.

Ein entscheidender Schritt wäre, eine Umgebung zu schaffen, die mehr Flexibilität ermöglicht. Forscher sollten zwischen Universitäten und Unternehmen hin- und herwechseln können, ohne ihre Karrierechancen zu gefährden. In den USA ist es selbstverständlich, dass Forscher bei Google oder Meta arbeiten und gleichzeitig auf internationalen Konferenzen publizieren. Solche Karrieremodelle müssen auch in Europa etabliert werden.

Zusätzlich sollten Unternehmen mehr Forschungslabore aufbauen, deren Forscherinnen und Forscher aktiv an internationalen Konferenzen teilnehmen und mit führenden Wissenschaftlern kooperieren. Derzeit fehlt es vielen europäischen Unternehmen an der Offenheit, solche Strukturen zu schaffen. Stattdessen dominieren Entwicklungsabteilungen, die meist intern arbeiten und kaum an der globalen wissenschaftlichen Diskussion teilnehmen. Die Forschungslabore der Unternehmen können auch an der

Universität angesiedelt sein, so wie in Amsterdam. Dort gibt es Forschungslabore von ca. zehn Personen, die von Unternehmen finanziert werden und völlig frei forschen können. Das ist ein sehr erfolgreiches Konzept, das erlaubte, dass Unternehmen Forscher ins Unternehmen holten, so die neuesten KI-Technologien mitbekamen und für sich nutzen konnten.

Ein Modell für Europa könnten KI-Institute sein, in denen Wissenschaftler und Industrie zusammenarbeiten. Diese Institute könnten als Schnittstelle fungieren, um neue Technologien in die Unternehmen zu bringen und gleichzeitig praxisnahe Forschung zu ermöglichen. Erste Ansätze gibt es bereits, etwa in Helsinki oder Tübingen, wo sogenannte »ELLIS Institute« eingerichtet wurden. Solche Institute können Forschern aus Universitäten und Unternehmen eine gemeinsame Plattform bieten und Orte sein, an dem verschiedene Akteure voneinander lernen. So könnten beispielsweise Entwickler aus Unternehmen von den neuesten akademischen Fortschritten profitieren, während Wissenschaftler praxisnahe Probleme aus der Industrie kennenlernen. Ein solches Modell würde nicht nur die Forschung beschleunigen, sondern auch dazu beitragen, innovative Technologien schneller in die Praxis zu bringen. Die Unternehmen könnten zudem direkt miteinander sprechen und sich über die Probleme bei der Einführung von KI-Technologie austauschen. Diese Probleme sind oft innere Widerstände, die fehlende Recheninfrastruktur oder die Hochskalierung von Methoden auf industrielle Anwendungen. Die Erfahrungen einer Firma könnten für andere Firmen allerdings sehr nützlich sein.

Ein weiteres Problem ist die Risikobereitschaft. In den USA gibt es eine Kultur des »Fail Fast«. Fehler zu machen, ist dort nicht das Ende, sondern ein notwendiger Schritt auf dem Weg zum Erfolg. In Europa, insbesondere in Deutschland, herrscht oft eine andere Mentalität. Hier geht es mehr darum, Risiken zu vermeiden und erst dann zu handeln, wenn alles perfekt ist. Das führt dazu, dass viele Projekte zu langsam oder gar nicht umgesetzt werden.

Nicht zuletzt fehlt in Europa das nötige Kapital. KI-Entwicklung ist teuer – besonders in der Hochskalierungsphase, wo riesige Rechenressourcen und hochqualifizierte Fachkräfte benötigt werden. Die großen Tech-Giganten in den USA und China haben nicht nur das Geld, sondern auch den politischen Rückhalt, um diese Summen zu investieren. Europäische Start-ups und Unternehmen hingegen müssen oft hart kämpfen, um überhaupt an halbwegs ausreichende Mittel zu kommen.

Trotz all dieser Hindernisse bin ich überzeugt, dass Europa jetzt eine zweite Chance hat. Wir befinden uns in einer Phase, in der es nicht mehr nur darum geht, riesige Modelle zu bauen. Es geht darum, diese Modelle auf spezifische Anwendungen und Branchen zuzuschneiden. Das ist eine Chance, bei der Europas Stärken – technisches Knowhow, spezialisierte Industriebranchen und eine exzellente Forschungstradition – ins Spiel kommen können.

Ein Beispiel dafür ist die Automobilindustrie. Selbstfahrende Autos sind nicht nur ein Marketing-Gag, sondern ein Paradebeispiel für die Anwendung von KI. Und hier könnte Europa tatsächlich wieder eine führende Rolle einnehmen – wenn wir es schaffen, die nötigen Technologien nicht nur zu entwickeln, sondern auch konsequent in die Praxis zu bringen. Ähnlich sieht es in anderen Branchen aus, etwa in der Pharmaindustrie, der Logistik oder im Maschinenbau.

Neues Denken

Der Schlüssel zur Industrialisierung der KI liegt in der Ausbildung einer neuen Generation von Fachkräften. Wir brauchen Ingenieure und Informatiker, die nicht nur die technischen Grundlagen verstehen, sondern auch das nötige Branchenwissen mitbringen, um KI in spezifischen Domänen anzuwenden. Das bedeutet, dass wir interdisziplinäre Ansätze fördern müssen. Ein Maschinenbauer sollte heute nicht mehr nur Maschinenbau verstehen, sondern auch Grundkenntnisse in KI haben.

Umgekehrt sollte ein KI-Forscher auch etwas von der Industrie verstehen, in der er arbeitet.

Wir brauchen eine neue Art des Denkens. Es reicht nicht, die bestehenden Prozesse nur ein bisschen effizienter zu machen. Wir müssen uns fragen: Wie können wir Dinge völlig neu denken? Wie können wir KI nutzen, um Innovationen zu schaffen, die vorher unmöglich waren? Hier ist Kreativität gefragt – und der Mut, auch mal unkonventionelle Wege zu gehen.

Ein weiteres großes Hindernis für die Verbreitung von KI in Europa sind die etablierten Strukturen in unseren traditionellen Industrien. Nehmen wir erneut die Automobilbranche oder die Pharmaindustrie. Viele Unternehmen sind hierarchisch organisiert, mit starren Abteilungen und Führungskräften, die oft seit Jahrzehnten in ihren Positionen sind. Wenn dann plötzlich eine neue Technologie wie KI auftaucht, löst das in solchen Strukturen häufig Angst aus. Es geht nicht nur um die Sorge, etwas nicht zu verstehen, sondern oft auch um die Angst, Kontrolle oder sogar den eigenen Job zu verlieren.

Ich habe das in meiner eigenen Arbeit oft erlebt. In der Pharmaindustrie zum Beispiel gab es enormen Widerstand gegen KI-gestützte Vorhersagen. Wir konnten mithilfe von Machine Learning relativ präzise sagen, welche Moleküle Nebenwirkungen verursachen könnten. Aber die Chemiker in den Unternehmen wollten das nicht akzeptieren. Ihre Haltung war: »Ich bin der Experte, ich verstehe die Chemie, und es kann doch nicht sein, dass eine Maschine etwas besser weiß als ich.« Es hat Jahre gedauert, diesen Widerstand aufzubrechen. Interessanterweise war es ein Generationswechsel, der letztendlich den Wandel gebracht hat. Als jüngere, technologieaffine Wissenschaftler und Führungskräfte in Schlüsselpositionen kamen, änderte sich die Einstellung zur KI radikal. Diese Generation hatte oft schon während ihres Studiums mit Machine Learning gearbeitet und verstanden, dass KI keine Bedrohung ist, sondern ein Werkzeug, das ihre Arbeit erleichtert und verbessert.

Besonders interessant finde ich, dass diese jungen Talente häufig gezielt experimentelle Gruppen in Unternehmen suchen – solche, die sich mit KI beschäftigen und neue Ansätze ausprobieren. Später übernehmen sie Führungspositionen und treiben den Wandel der Unternehmenskultur weiter voran.

In der Pharmaindustrie begann dieser Wandel der Unternehmenskultur etwa ein Jahrzehnt nach den ersten Anwendungen von KI. 2006 haben wir mit unseren Vorhersagesystemen angefangen, und bis 2015 waren es oft noch isolierte Gruppen innerhalb der Unternehmen, die sich mit diesen neuen Technologien beschäftigten. Erst als die Mitglieder dieser Gruppen in Führungspositionen aufstiegen, wurde KI breiter akzeptiert. Heute ist es in der Medikamentenentwicklung unvorstellbar, ohne KI zu arbeiten, vor allem in der frühen Phase, wo es darum geht, potenzielle Nebenwirkungen vorherzusagen.

Ein Beispiel, das mir besonders in Erinnerung geblieben ist, war eine Situation in der medizinischen Forschung. Wir hatten Gene priorisiert, die für eine bestimmte Krankheit relevant sein könnten, und unsere Rangliste den Medizinern vorgelegt. Aber sie nahmen nicht die am höchsten priorisierten Gene, sondern griffen sich ein Gen aus der Mitte heraus – eines, das sie aus ihrer eigenen Erfahrung kannten. Das hat uns frustriert, weil wir wussten, dass unsere Rangliste auf soliden Daten basierte. Später stellte sich heraus, dass sie unsere Vorschläge nur dann akzeptierten, wenn diese mit ihrem eigenen Wissen übereinstimmten. Sie waren sehr resistent gegen neues Wissen.

Das war ein klares Beispiel für den Widerstand gegen die »Black Box«-Natur von KI. Viele traditionelle Branchen tun sich schwer damit, auf ein System zu vertrauen, dessen Mechanismen sie nicht vollständig verstehen.

Ein Problem ist, dass viele Führungskräfte schlicht nicht wissen, wie KI funktioniert. Sie unter- oder überschätzen die Möglichkeiten der KI. Einerseits sehen sie einfach nicht, wo und wie KI in ihren Unternehmen eingesetzt werden kann. Andererseits überschätzen sie die Fähigkeiten der KI und sind dann enttäuscht, dass es nicht funktioniert hat. Bei KI denken

viele Führungskräfte an etwas Undurchschaubares, das sie nicht kontrollieren können. Hier braucht es dringend Schulungen und Aufklärungsarbeit, um den Zugang zu erleichtern.

Kultureller Wandel

Trotz all dieser Hindernisse sehe ich einen Hoffnungsschimmer: die neue Generation von Fachkräften. Junge Ingenieure, Chemiker und Informatiker wachsen mit einem völlig anderen Verständnis von Technologie auf. Für sie ist KI kein Fremdkörper, sondern ein integraler Bestandteil ihres Werkzeugsatzes. Sie sehen KI nicht als Konkurrenz, sondern als Ergänzung ihrer eigenen Fähigkeiten. Das ist ein grundlegender kultureller Wandel, der Europa helfen könnte, den Rückstand aufzuholen. Diese Generation ist bereit, Risiken einzugehen, neue Wege zu beschreiten und traditionelle Strukturen zu hinterfragen. Sie könnte der Schlüssel sein, um die Potenziale der KI in unseren traditionellen Industrien endlich voll auszuschöpfen.

Die Adaption von KI ist kein technisches Problem, sondern vor allem eine Frage von Kultur, Bildung und Führung. Wenn wir es schaffen, die Vorurteile und Ängste gegenüber KI abzubauen und gleichzeitig die neue Generation zu fördern, könnten wir in Europa einen echten Wandel erleben. Der Erfolg hängt davon ab, wie schnell und wie umfassend wir diesen kulturellen Wandel umsetzen können. Wir müssen die KI nicht nur in unsere Unternehmen bringen, sondern auch in unsere Köpfe. Denn nur, wenn wir verstehen, was KI ist und was sie leisten kann, werden wir sie als das nutzen, was sie ist: ein Werkzeug, das uns nicht ersetzt, sondern unsere Möglichkeiten erweitert.

Europa hat im Bereich der KI-Forschung viele talentierte Wissenschaftler und exzellente Universitäten. Doch der Übergang von der Grundlagenforschung zur Kommerzialisierung bleibt eine Schwäche. Während Unternehmen wie Google und OpenAI ihre Modelle schnell

produktiv machen, fehlt in Europa oft die Infrastruktur und die Verknüpfung zwischen Wissenschaft und Industrie.

Ich bin überzeugt, dass Europa durch gezielte Investitionen in Infrastruktur – sogenannte »Moonshot Facilities« – aufholen kann. Solche Einrichtungen könnten es ermöglichen, auch in Europa großskalige Modelle wie ChatGPT zu entwickeln. Doch die bürokratischen Hürden und der Wettbewerb um Ressourcen erschweren diesen Fortschritt. In meinen Gesprächen mit europäischen Politikern habe ich immer wieder betont, wie wichtig es ist, schnell und unbürokratisch zu handeln, um nicht den Anschluss zu verlieren.

Zu viel Regulierung?

Ich war von Beginn an intensiv in Diskussionen zur KI-Regulierung involviert, vor allem durch meine Arbeit mit ELLIS (European Laboratory for Learning and Intelligent Systems; ich bin dort Schatzmeister) und den Austausch mit EU-Parlamentariern. Der »AI Act« der EU, eine gesetzliche Regelung, die sicherstellen soll, dass KI-Systeme in Europa ethisch vertretbar und sicher eingesetzt werden, hat von Anfang an eine Menge Debatten ausgelöst. Leider lief der erste Entwurf des Gesetzes komplett in die falsche Richtung. Es war den Mitgliedstaaten und ihrer jeweiligen Ratspräsidentschaft – etwa den Finnen – zu verdanken, dass der Entwurf in eine vernünftigere Richtung korrigiert wurde. Dennoch gibt es nach wie vor wesentliche Schwächen.

Die Frage, ob wir eine Regulierung für KI brauchen, beantworte ich klar mit »Ja«. Viele Unternehmen und Entwickler fragen sich: Was ist erlaubt, was nicht? Gerade wenn KI in Anwendungen integriert wird, ist es essenziell, klare Rahmenbedingungen zu schaffen. Die Idee, KI in Risikostufen einzuteilen, finde ich grundsätzlich gut. Doch die Umsetzung bleibt problematisch, vor allem bei Definitionen wie »generative KI« oder »General Purpose AI«.

Zum Beispiel ist die Definition von »KI« im AI Act äußerst fragwürdig. Folgt man der EU-Definition, gilt schon eine einfache mathematische Funktion als KI. Damit wird der Begriff KI so weit gedehnt, dass er faktisch alles umfasst. Wir Wissenschaftler haben über Jahrzehnte viel Zeit investiert, eine sinnvolle Definition von KI zu erarbeiten. Die aktuelle Regelung ignoriert diese Arbeit.

Ein gutes Beispiel für eine fundierte Definition ist die von Edward Feigenbaum, einem Pionier auf dem Gebiet der künstlichen Intelligenz. Feigenbaum definiert KI nicht einfach als eine Funktion, die Eingabedaten verarbeitet und einen Output liefert. Nach Barr und Feigenbaum (1981) ist KI eine Maschine, die kognitive Fähigkeiten wie der Mensch aufweist, z. B. zu lernen, zu planen, logisch zu folgern, Probleme zu lösen und neue Fähigkeiten zu erwerben. Der letzte Punkt »neue Fähigkeiten zu erwerben« wurde von mir hinzugefügt, da François Chollet (ein bekannter KI-Forscher) sagt: »Intelligenz ist die Fähigkeit, neue Fähigkeiten zu erwerben«. Solche jahrzehntelang erprobten Definitionen fehlen im AI Act völlig, was die Diskussion unnötig erschwert.

Besonders kritisch sehe ich die Klassifizierung von »General Purpose AI« (GPAI), also KI-Systemen mit allgemeinem Anwendungszweck, die potenziell ein systemisches Risiko darstellen können. Die EU hat hier keine klare Definition. Es gibt lediglich eine technische Grenze: Modelle, die mit einer Rechenleistung von mehr als 10^{25} FLOPS (Floating Point Operations) trainiert wurden, fallen darunter. Das Problem dabei ist, dass diese Schwelle willkürlich erscheint und mit der Weiterentwicklung der Technologie angepasst werden müsste.

Zusätzlich fordert der AI Act umfangreiche Risikoanalysen und eine strikte Kontrolle der Datenquellen für diese Art von KI. Doch was bedeutet das in der Praxis? Man weiß ja nicht, ob das eigene System darunterfällt oder nicht. Aber man muss schon zu Anfang Daten und Risikoanalysen machen, obwohl man nicht mal weiß, wie groß man das Modell machen muss, um eine bestimmte Leistung zu bekommen. Diese Unsicherheit

erschwert die Entwicklung und bringt europäische Firmen in eine schwierige Lage. Unternehmen wie Google oder Meta, die ihre KI-Systeme bereits weltweit etabliert haben, könnten diese Anforderungen einfach ignorieren. Europäische Start-ups hingegen riskieren, durch strikte Auflagen benachteiligt zu werden. Das führt zu einer Situation, in der große US-Unternehmen ihre Systeme verteidigen können, während kleinere europäische Unternehmen das Nachsehen haben.

Diese Problematik spiegelt sich auch in der Frage nach Open-Source-Software wider. Während Open-Source-Ansätze in vielen Bereichen vom AI Act ausgenommen sind, gilt dies nicht für General Purpose AI. Selbst wenn ein Modell öffentlich zugänglich ist, könnten Vorschriften dennoch verhindern, dass es genutzt wird, was die Innovationskraft weiter hemmt.

Diese Unsicherheit erinnert mich stark an die Einführung der DSGVO (Datenschutzgrundverordnung). In einem Projekt mit einem Automobilunternehmen ging es beispielsweise um die Auswertung von Videodaten, die für die Entwicklung selbstfahrender Autos verwendet werden sollten. Das Projekt wurde fast ein Jahr lang auf Eis gelegt, weil unklar war, ob die Daten überhaupt genutzt werden dürfen. Am Ende stellte sich heraus, dass die Nutzung rechtlich zulässig war, da es um allgemeine Forschung und nicht um die Verarbeitung personenbezogener Daten ging. Doch die Rechtsunsicherheit hatte bereits erheblichen Schaden angerichtet. Ähnliches droht nun mit dem AI Act. Entwickler könnten Projekte nicht starten oder Unternehmen gar nicht erst gründen, weil sie nicht wissen, ob ihre Anwendungen den Regularien entsprechen.

Ein weiteres Problem ist die Wettbewerbsverzerrung. Firmen, die bereits auf dem Markt etabliert sind, können ihre Position nutzen, um neue Akteure zu blockieren. Große US-Unternehmen könnten beispielsweise drohen, ihre Produkte vom Markt zu nehmen, wenn ihre Systeme den neuen Vorschriften nicht entsprechen. Europäische Unternehmen, die diese Produkte nutzen, hätten dann keine Alternativen. Das Ergebnis wäre eine weitere Stärkung der großen Akteure und eine Schwächung der europäischen Innovationslandschaft.

Europa hat das Potenzial, im Bereich KI weltweit führend zu sein. Doch dafür müssen wir die Innovationsbereitschaft der Industrie stärken. Aus meiner Sicht sind hier zwei Faktoren entscheidend: eine engere Verzahnung von Industrie und Wissenschaft sowie ein Wandel in der Denkweise. Das heißt, wir benötigen die Etablierung einer neuen Kultur.

In vielen traditionellen Branchen herrscht noch immer eine gewisse Skepsis gegenüber neuen Technologien. Doch genau wie in der Pharmaindustrie vor einigen Jahren wird es auch in anderen Branchen darauf ankommen, dass eine neue Generation von Fachkräften die Technologie akzeptiert und integriert. Das könnte durch gezielte Bildungsprogramme und die Förderung von Start-ups beschleunigt werden.

Der Staat spielt ebenfalls eine zentrale Rolle in diesem Prozess. Einerseits muss er gezielt in die Forschung investieren, andererseits sollte er Anreize schaffen, damit Unternehmen und Universitäten enger zusammenarbeiten.

Die Idee eines »CERN für KI« führt in die falsche Richtung. KI braucht keinen zentralen Ort für die Forschung, denn KI-Forscher arbeiten dezentral. Sie treffen sich bei den großen KI-Konferenzen und müssen nicht noch zusätzlich einen Ort zum Zusammentreffen haben. Anders ist es jedoch bei Recheninfrastrukturen. Diese müssen zentral sein, da Nodes mit GPUs mit schnellen Verbindungen ausgestattet werden müssen. Wir haben zum Beispiel xLSTM auf 32 Nodes à 8 GPUs, also auf 256 GPUs, trainiert, die miteinander mit schnellen Verbindungen vernetzt waren. Solche Rechenzentren sollten jedoch dort stehen, wo Energie und Kühlung billig sind, etwa auf Island. Wie gesagt, wären aber auch Orte, wo sich Industrie, Start-ups und Wissenschaft treffen, wünschenswert.

Die Politik spielt eine Schlüsselrolle, wenn es darum geht, Rahmenbedingungen für Innovation zu schaffen. Dabei geht es nicht nur um den AI Act, sondern auch um Förderprogramme, die den Einstieg für kleine Unternehmen erleichtern. Es braucht klare, praxisnahe Regularien, die Innovation fördern, statt sie durch bürokratische Hürden zu blockieren. Die kommende Dekade wird entscheidend sein, um Europas Position im

globalen Wettbewerb zu sichern. Mit KI haben wir die Chance, traditionelle Industrien wie die Auto- und Pharmaindustrie zu revolutionieren und gleichzeitig neue Märkte zu erschließen. Doch dazu müssen wir mutiger werden. Die Innovationskraft Europas liegt in der Fähigkeit, Wissenschaft und Wirtschaft zu verbinden. Wenn wir diese Stärke nutzen, können wir nicht nur mithalten, sondern den Weg in eine technologische Zukunft aktiv gestalten.

Europa steht vor einer Herausforderung, die ich in Gesprächen immer wieder anspreche: Wir hängen zu stark von den Vereinigten Staaten ab, wenn es um KI-Technologie geht. Die großen KI-Modelle, die Werkzeuge, die Infrastruktur – all das stammt überwiegend aus den USA. Die meisten europäischen Firmen arbeiten auf der Amazon-Cloud AWS oder bei Oracle, also auf amerikanischer Recheninfrastruktur. Dabei stellt sich aber nicht nur eine wirtschaftliche Frage, sondern auch eine politische. Was passiert, wenn die Amerikaner eines Tages beschließen, uns den Zugang zu diesen Technologien zu erschweren oder sogar ganz zu blockieren? Was, wenn sie den »Hahn zudrehen«? Diese Frage ist keine bloße Spekulation. Wir haben in der Vergangenheit gesehen, wie geopolitische Spannungen technologische Kooperationen beeinträchtigen können.

Ein weiterer Faktor, der Europa in eine schwierige Position bringt, ist der bereits erwähnte AI Act. Ich bin absolut für ethische Standards in der KI. Aber wir müssen uns bewusst sein, dass diese Standards auch Nachteile haben können. Während europäische Firmen gezwungen sind, strenge Vorschriften einzuhalten, haben amerikanische Unternehmen oft freie Hand. Das schafft ein ungleiches Spielfeld. Wenn die USA ihre Technologien ohne Rücksicht auf europäische Regulierungen weiterentwickeln, könnten sie einen massiven Vorsprung aufbauen. Europäische Firmen hingegen müssen sich mit bürokratischen Hürden auseinandersetzen, die ihre Innovationskraft bremsen.

Man spricht oft von sogenannten »Moonshot-Projekten«, bei denen Europa sich ambitionierte Ziele wie die Bekämpfung von Krebs setzen könnte. Aber meiner Meinung nach ist das nicht der richtige Ansatz. Was

wir brauchen, sind »Moonshot Facilities« – leistungsstarke Infrastrukturen, die es uns überhaupt erst ermöglichen, solche Projekte durchzuführen. Ohne die notwendigen Rechenressourcen können wir nicht einmal anfangen, ein europäisches Pendant zu ChatGPT zu entwickeln, geschweige denn andere große KI-Projekte umzusetzen.

Ohne diese Basis können wir nicht wirklich wettbewerbsfähig werden. KI-Projekte, die weltweit Aufsehen erregen, wie AlphaGo, AlphaFold oder ChatGPT, sind nur möglich, weil es in den USA Unternehmen gibt, die Zugang zu Zehntausenden GPUs haben. Diese Ressourcen stehen Forschern dort quasi auf Knopfdruck zur Verfügung. Derzeit haben wir in Europa zwar einige leistungsstarke Rechenzentren wie Leonardo in Italien, Marenostrum5 in Spanien oder Lumi in Finnland, aber die Kapazitäten sind stark fragmentiert. Forscher müssen sich durch komplizierte Bewerbungsverfahren kämpfen, um Zugang zu kleinen Teilen dieser Ressourcen zu erhalten. Solche Prozesse dauern oft Monate, und am Ende bekommt man nur einen Bruchteil der Rechenleistung, die man eigentlich benötigt.

Das ist ein riesiges Problem, wenn wir mit der Geschwindigkeit der Amerikaner mithalten wollen. Dort arbeiten Unternehmen wie Google oder Meta mit Tausenden von GPUs, die ihren Forschern sofort zur Verfügung stehen. Ich erinnere mich an Studenten, die von einem Praktikum bei Google zurückkamen und berichteten, dass sie über Nacht auf 1.024 GPUs gerechnet haben. In Europa dagegen kämpfen wir darum, überhaupt 10 GPUs für ein Projekt zu bekommen.

Was wir brauchen, ist eine leistungsstarke, flexible und zentralisierte Infrastruktur, die für Forscher leicht zugänglich ist. Wir brauchen GPU-Zentren, die so ausgelegt sind, dass Forscher auf Tausende von GPUs gleichzeitig zugreifen können, und zwar ohne langwierige Antragsverfahren. Wichtig ist, dass diese Ressourcen flexibel und schnell verfügbar sind. Es ist frustrierend zu sehen, wie viel Zeit und Ressourcen wir in bürokratischen Prozessen verlieren, während andere Regionen voranschreiten.

Ein gutes Beispiel ist der Versuch, ein europäisches ChatGPT zu entwickeln. Aktuell ist das schlichtweg unmöglich, weil wir nicht die notwendige Rechenleistung haben. Um ein solches Modell von Grund auf zu trainieren, benötigt man Zehntausende GPUs über mehrere Monate hinweg. Das kostet Hunderte Millionen Euro, und solche Ressourcen sind in Europa einfach nicht vorhanden. Selbst in den großen europäischen Rechenzentren sind die Kapazitäten begrenzt und stark verteilt. Das macht es unmöglich, solche Projekte effizient umzusetzen. Auch ein schnelles Modell wie das von DeepSeek (DeepSeek v3) hat zuerst Zehntausende von GPUs gebraucht und die Methoden zu entwickeln, die das Training und die Inferenz des Modells schnell gemacht haben. Andererseits sieht man dadurch aber, dass man mit guten Ideen auch Rechenleistung einsparen kann. Zurzeit geht die Entwicklung eher wieder in Richtung kleinerer, effizienter Modelle, wie Low Precision (DeepSeek nimmt 8 Bits zur Darstellung von Zahlen statt 32 oder 16 Bits), Mixture of Experts (es rechnet immer nur ein Expertenmodell, statt das ganze Modell) oder Sparse Matrices (Matrizen sind spärlich besetzt).

Ich habe wiederholt darauf hingewiesen, dass wir dringend in eine Infrastruktur investieren müssen. Die Kosten wären zwar hoch – möglicherweise mehrere Milliarden Euro –, aber sie wären eine Investition in die Zukunft. Denn ohne diese Ressourcen können wir keine konkurrenzfähigen Modelle entwickeln, und Europa wird weiter an Boden verlieren. Nun endlich wurden solche Projekte von der Präsidentin der Europäischen Kommission angekündigt.

Ein weiteres großes Problem ist die Bürokratie. Selbst wenn wir die notwendigen Ressourcen hätten, dauert es oft viel zu lange, bis Projekte genehmigt werden. Das ist in der heutigen Zeit, in der sich die KI-Forschung rasant entwickelt, ein enormes Hindernis. In den USA können Unternehmen innerhalb weniger Tage Entscheidungen treffen und Projekte starten. In Europa dagegen dauert es Monate, bis ein Antrag genehmigt wird.

Ich habe vorgeschlagen, ein unabhängiges wissenschaftliches Gremium einzurichten, das schnell und unbürokratisch über die Vergabe von Ressourcen entscheidet. So könnten Projekte wie ein europäisches ChatGPT zeitnah umgesetzt werden, ohne dass sich Forscher durch endlose Verwaltungsprozesse kämpfen müssen.

Was mich frustriert, ist, dass viele Politiker nicht verstehen, wie dringend dieses Problem ist. Ich habe mehrfach versucht, die Bedeutung solcher Infrastrukturen zu erklären, aber oft stoßen solche Vorschläge auf taube Ohren. Es scheint, als würde man lieber kleine Projekte finanzieren, statt die Grundlagen zu schaffen, die wir wirklich brauchen. Das hat sich mit der Ankündigung der Präsidentin der Europäischen Kommission, Ursula von der Leyen, geändert. Sie hat zusätzliche 20 Milliarden Euro aus dem EU-Haushalt für die Entwicklung von künstlicher Intelligenz (KI) in Europa zugesagt. Weitere 150 Milliarden sollen aus der Privatwirtschaft und von Investoren kommen. Es wird von 200 Milliarden Euro für den KI-Ausbau in Europa gesprochen. Das hört sich toll an – ich hoffe, die Umsetzung ist auch toll.

Ein weiteres Hindernis ist die politische Dimension. Großprojekte in Europa scheitern oft daran, dass sich die Mitgliedstaaten nicht einigen können, wo solche Einrichtungen entstehen sollen. Jeder will seinen Anteil, und am Ende passiert gar nichts. Wir müssen endlich über diese regionalen Interessen hinwegsehen und uns auf das große Ganze konzentrieren. In Gesprächen mit Entscheidungsträgern, sowohl in Deutschland als auch auf europäischer Ebene, habe ich immer wieder darauf hingewiesen, dass wir ohne solche Infrastrukturen keine Chance haben, mit den USA und China mitzuhalten.

Nun haben sie endlich hingehört. Die Zeit drängt. KI entwickelt sich mit einer Geschwindigkeit, die wir vor ein paar Jahren noch für unmöglich gehalten hätten. Wenn wir jetzt nicht investieren, wird Europa in der globalen KI-Landschaft nur noch eine Randrolle spielen. Das können wir uns nicht leisten – weder wirtschaftlich noch gesellschaftlich. Deshalb ist es gut, dass nun so viel Geld für die KI in die Hand genommen wird.

Die Frage, welche Bereiche Europa priorisieren sollte, ist nicht leicht zu beantworten. Ich persönlich glaube, dass wir ein wissenschaftliches Gremium brauchen, das diese Entscheidungen trifft – ein Gremium, das unabhängig, schnell und auf Basis fundierter Daten agiert.

Wie bereits erwähnt, liegt die größte Chance Europas meiner Meinung nach in der dritten Phase der KI-Entwicklung: der Industrialisierung von KI. Hier haben wir noch eine echte Chance, weil wir in Europa über zahlreiche hochspezialisierte Branchen verfügen – sei es die Automobilindustrie, der Maschinenbau oder die Pharmaindustrie. Wenn wir es schaffen, KI in diesen Branchen effektiv einzusetzen, könnten wir wieder eine Führungsrolle übernehmen.

10

KI kontrollieren

Viele Menschen begegnen KI mit Skepsis. Dafür gibt es mehrere Ursachen. Einerseits prägen Filme wie *Terminator* oder *The Matrix* eine Vorstellung von KI als Bedrohung für die Menschheit – ein Hollywood-Narrativ, das tiefe Spuren im kollektiven Bewusstsein hinterlassen hat. Andererseits ist die KI sehr komplex und nur schwer verständlich, was häufig zu einem Gefühl des Kontrollverlustes und zu Misstrauen führt. Besonders Technologien, deren Funktionsweise nicht transparent ist, wecken schnell Zweifel. Hinzu kommt eine Art psychologischer Widerstand: Die Vorstellung, dass uns Maschinen bald in vielen kognitiven Fähigkeiten überlegen sein könnten, erschüttert unser Selbstverständnis und erzeugt Unbehagen, ja Angst. KI ist die letzte Kränkung der Menschheit, da Maschinen uns nun in unserer letzten Domäne, in der wir ihnen überlegen waren, auch bald voraus sein könnten.

Viele Befürchtungen halte ich für maßlos übertrieben. KI wird die Menschheit nicht auslöschen. Sie hat keinen eigenen Willen und sie verfolgt keine Interessen. Die Angst, dass KI-Systeme plötzlich emergente Fähigkeiten entwickeln – also Fähigkeiten, die über das hinausgehen, was sie ursprünglich gelernt haben –, ist unbegründet. Wir Menschen neigen dazu, in KI mehr zu sehen, als tatsächlich vorhanden ist. Wenn ein Modell eine beeindruckende Antwort liefert, interpretieren wir oft Intelligenz oder Absicht hinein, wo lediglich gelernte Texte abgespult und mit statistischen Mustern und Wahrscheinlichkeiten kombiniert werden. Diese Projektionen können dazu führen, dass wir KI-Systemen Fähigkeiten zuschreiben, die sie nicht besitzen, was wiederum die Erwartungen und Ängste rund um KI verstärken kann.

Was wirklich passiert, ist, dass die Modelle Muster in den Daten erkennen und diese effizient kombinieren. Das mag überraschend wirken, doch es ist letztlich eine Frage von Skalierung und Optimierung von Datenverarbeitung, nicht von echtem Verständnis oder Bewusstsein. Wir sollten daher keine Science-Fiction-Vorstellungen auf die KI projizieren. Solche apokalyptischen Szenarien lenken nur von den wirklichen Gefahren ab.

Kontrolle behalten

Für mich gibt es klare Grenzen, die wir in der KI-Forschung nicht überschreiten sollten. Besonders kritisch sehe ich die Vorstellung einer Art unkontrollierter »KI-Evolution«, in der die besten Systeme überleben, ohne dass wir als Menschen ihre Entwicklung lenken. Dann würde die KI auch in eine Evolution einsteigen und um Ressourcen kämpfen müssen. Zurzeit bestimmen aber Menschen, welche KI gut funktioniert hat und weiterentwickelt wird und welche KI nicht so gut war und gelöscht wird. Solange wir Menschen die Evolution der KI bestimmen, können sich keine ungewünschten KIs entwickeln.

Wir müssen auch die Kontrolle darüber behalten, welche Entscheidungen diese KI-Systeme treffen dürfen und welche nicht. Es ist essenziell, dass wir als Menschen immer die Verantwortung tragen und nicht zulassen, dass KI-Systeme uns überholen – nicht in ihrer Leistungsfähigkeit, sondern in ihrer Unabhängigkeit.

Eines der größten Risiken sehe ich in der Entwicklung von KI für Waffensysteme. Die Vorstellung, dass autonome Systeme darüber entscheiden könnten, ob ein Mensch lebt oder stirbt, widerspricht allen ethischen Prinzipien, die ich vertrete. Ähnlich sehe ich den Einsatz von KI zur Unterstützung diktatorischer Regime. Solche Technologien können genutzt werden, um Überwachungssysteme zu perfektionieren oder Propaganda effizienter zu gestalten – ein Missbrauch, der zutiefst

beunruhigend ist. Es ist mir jedoch klar, dass KI in Rüstung und Krieg eingesetzt werden wird – so wie bisher jede andere Technologie, wie zum Beispiel Autos oder das Telefon. Aber KI-Systeme speziell für den Krieg zu entwickeln, sehe ich als ethisch nicht vertretbar.

Ein Thema, das mich immer wieder beschäftigt, sind humanoide Roboter. Viele Unternehmen und Forschungsprojekte scheinen davon fasziniert zu sein, Maschinen zu entwickeln, die wie Menschen aussehen und sich möglichst menschenähnlich verhalten. Ich halte das für einen Irrweg. Sobald ein Roboter humanoid gestaltet ist, neigen wir Menschen dazu, ihm menschliche Eigenschaften zuzuschreiben. Wir interpretieren Emotionen in Bewegungen oder Mimik hinein, die gar nicht vorhanden sind. Das führt zu Missverständnissen und kann gefährlich werden, wenn wir diesen Maschinen mehr zutrauen, als sie tatsächlich können. Natürlich gibt es spezialisierte Roboter, die nützlich sein können – etwa solche, die Böden reinigen oder in der Altenpflege einfache Aufgaben übernehmen. Aber einen universellen, menschenähnlichen Roboter zu bauen, der alles kann, halte ich für unrealistisch und kontraproduktiv. Ein solcher Roboter müsste die Welt in ihrer ganzen Komplexität verstehen – ein Ziel, von dem wir meiner Meinung nach noch weit entfernt sind.

Eine zentrale Frage ist, wie sich KI auf die Arbeitsplätze in der Wirtschaft auswirken wird. Meine Beobachtung ist, dass Arbeitsplätze durch neue Technologien nicht verschwinden, sondern sich verändern. Denken wir an die Einführung des Computers: Anstatt Menschen zu ersetzen, hat er neue Berufsfelder geschaffen. Doch die Art der Arbeit hat sich dabei stark gewandelt. Mit der KI werden monotone und repetitive Aufgaben zunehmend automatisiert. Das bedeutet nicht, dass die Jobs verschwinden, sondern dass sie interessanter und abwechslungsreicher werden können. Ein Beispiel: Ein Autodesigner kann heute mithilfe von KI sofort simulieren, wie sich seine Entwürfe im Windkanal verhalten. Dadurch wird sein kreativer Prozess effizienter und spannender, weil er sofort Feedback erhält und neue Ansätze ausprobieren kann.

Ähnliches sehen wir bereits beim Programmieren. Mit KIgestützten Tools wie Code-Co-Piloten bekommen Entwickler Vorschläge und Templates, die viel Vorarbeit leisten. Statt sich um die grundsätzliche Programmstruktur zu kümmern, können Programmierer ihre Energie darauf verwenden, innovative Lösungen in die Programmstruktur einzubauen.

Natürlich bringt diese Unterstützung auch Herausforderungen mit sich. Eine häufige Sorge ist, dass wir durch KI »verblöden« könnten, da sie uns Denk- und Lernprozesse abnimmt. Das ist nicht von der Hand zu weisen. Rechtschreibung ist ein gutes Beispiel: Heute korrigieren Textverarbeitungsprogramme automatisch unsere Fehler. Es stellt sich die Frage, welche Fähigkeiten wir wirklich brauchen und zu welchem Grad wir sie erlernen sollen. Wir benötigen bestimmte Fähigkeiten, die zwar die KI auch hat, aber die wichtig sind, um die KI optimal einsetzen zu können.

In Zukunft werden Menschen und KI verstärkt Hand in Hand arbeiten. Die KI wird nicht nur ein Werkzeug sein, sondern ein Partner, der uns hilft, effizienter und innovativer zu sein. Die Herausforderung wird darin bestehen, diese Zusammenarbeit so zu gestalten, dass sie für alle Beteiligten – sowohl für Unternehmen als auch für die Arbeitnehmer – von Vorteil ist. Menschliche Kreativität, Problemlösung und kritisches Denken werden dabei weiterhin wichtig bleiben.

Vertrauen

Neuronale Netze sind eine »Black Box«, wir verstehen nicht genau, wie sie zu ihren Entscheidungen kommen. Ein Thema, das mich immer wieder beschäftigt, ist der Konflikt zwischen Erklärbarkeit und Performance. Wenn ein Modell besser als ein Mensch sein soll, dann kann es nicht immer vollständig erklärbar sein. Warum? Weil die besten Modelle oft Dinge tun, die so komplex sind, dass sie für uns Menschen schwer nachzuvollziehen sind. Wenn ein Modell alles so einfach erklären könnte, dass ein Mensch es versteht, dann wäre es wahrscheinlich nicht besser als der Mensch selbst.

Das führt besonders in Bereichen wie der Medizin zu einem Dilemma: Ärzte und Patienten wollen nachvollziehen können, warum ein Modell eine bestimmte Diagnose stellt. Gleichzeitig verlangen wir, dass diese Modelle bessere Ergebnisse liefern als menschliche Experten. Das bedeutet, wir müssen akzeptieren, dass nicht alles vollständig erklärbar ist. Ein möglicher Kompromiss ist, dass das Modell die Hauptfaktoren seiner Entscheidung erklärt und komplexere Zusammenhänge im Hintergrund bleiben.

In der Medizin sehe ich ein großes Potenzial für KI, aber auch besondere Herausforderungen. Vertrauen ist hier das A und O. Ein Modell muss zeigen können, warum es zu einer bestimmten Diagnose kommt. Gleichzeitig darf es nicht an Leistung einbüßen. Ein Ansatz, den ich befürworte, ist die sogenannte »Teil-Erklärbarkeit«. Das Modell könnte sagen: »Die Diagnose basiert auf einem hohen Blutdruck und einer auffälligen Blutanalyse.« Das sind für Ärzte nachvollziehbare Faktoren. Andere, komplexere Muster könnten als unterstützende Hinweise angegeben werden, ohne dass sie den Arzt überfordern. Die heutigen LLMs geben oft eine Chain of Thought neben der Lösung aus, das heißt, sie liefern die Zwischenschritte, die zu einer Lösung notwendig waren. Dadurch versteht man die Lösung besser und kann sie auch verifizieren. Die Idee der Chain of Thought wird immer häufiger verwendet, auch wenn es sich nicht um mathematische Probleme handelt. Hier sind Prompts wie »Let's think step by step« sehr hilfreich, um Zwischenschritte und Zwischenergebnisse neben der Lösung zu erhalten. Das gibt viel mehr Einsicht in die Lösung. Methoden wie DeepSeek R1 verwenden »Monte Carlo Tree Search« zum Durchlaufen solcher Zwischenschritte. Somit können diese dann auch einem Benutzer zur Verfügung gestellt werden und die Antwort wird besser verständlich.

Wir Menschen verstehen die Lernprozesse in kleinen, einfachen Systemen recht gut. Doch bei den großen Modellen wird es schwierig. Hochskalierte Systeme haben eine eigene Dynamik, die wir oft nicht nachvollziehen können. Manchmal funktionieren sie hervorragend, dann wieder gar nicht – und niemand weiß genau, warum. Wenn während des

Lernens etwas schiefgeht, kann das Modell plötzlich instabil werden oder nicht mehr lernen. Ich würde diese Prozesse gern besser verstehen – und herausfinden, was stabile und erfolgreiche Lernmethoden ausmacht.

Es bleibt die Frage, ob wir immer verstehen müssen, wie die KI zu ihren Ergebnissen kommt. Für mich ist das keine Voraussetzung. Es ist okay, wenn wir nicht alle Details nachvollziehen können, solange die Ergebnisse konsistent und überprüfbar sind. Manchmal ist es eben genug, wenn wir die Antworten nutzen können – auch wenn wir nicht den gesamten Prozess dahinter verstehen.

Es ist kein neuer Gedanke, dass wir Werkzeuge nutzen, deren Funktionsweise wir nicht vollständig verstehen. Niemand versteht alle Details eines Handys, doch wir vertrauen darauf, dass es funktioniert. Ein weiteres Beispiel ist die Genetik. Wenn wir die Expression von Hunderten Genen betrachten, können wir keine einfache Regel ableiten, die uns erklärt, was diese Muster bedeuten. Es fehlt uns an Konzepten, um diese Komplexität zu greifen. Eine KI könnte solche Muster analysieren und uns Ergebnisse liefern, die für uns unverständlich sind, da wir nicht so viele Komponenten gleichzeitig kombinieren können.

Das Schachspiel ist ein gutes Beispiel dafür, wie KI menschliche Fähigkeiten übertreffen kann. Eine KI wie AlphaZero spielt Schach auf einem Niveau, das für uns unerreichbar ist. Schach ist eines meiner liebsten Beispiele dafür, wie KI sinnvoll eingesetzt werden kann. Natürlich wissen wir, dass kein Mensch mehr eine Chance gegen die besten Schachcomputer hat. Und doch spielen wir weiter – nicht, um zu gewinnen, sondern um zu lernen und Freude daran zu haben. Ich spiele selbst gelegentlich gegen Computer. Später analysiere ich dann meine eigenen Fehler. Was hätte ich besser machen können? War mein Angriff strategisch klug oder überhastet? Die KI zeigt mir Wege, die ich übersehen habe. Sie erweitert meinen Horizont und bereichert mein Spiel. Aber der wahre Reiz beim Schach bleibt der menschliche Gegner – die Emotionen, die Dynamik, die Unvorhersehbarkeit, die mit ihm einhergehen.

Wie beim Schach sollte es unser Ziel sein, KI als Werkzeug zu nutzen, nicht als Ersatz für etwas oder jemanden. Wir können von ihr lernen, aber wir sollten uns ihr nicht unterwerfen. KI kann uns helfen, klüger und besser zu werden, ohne uns unsere Eigenständigkeit zu nehmen.

Aktuelle KI-Modelle wie ChatGPT sind nützlich für eine Vielzahl von Aufgaben, insbesondere bei der schnellen Verarbeitung von Informationen. Zum Beispiel können sie helfen, Texte zu verbessern, Informationen effizient zu recherchieren oder allgemeine Fragen zu beantworten. Ich sehe diese Anwendungen als unproblematisch, solange es um standardisierte oder unkritische Aufgaben geht. Doch wenn es um wichtige Entscheidungen geht, etwa finanzielle Investitionen oder medizinische Diagnosen, würde ich mich niemals ausschließlich auf ein KI-System verlassen. In solchen Fällen ist ein menschlicher »Quercheck« entscheidend, um Plausibilität und mögliche Risiken zu bewerten. Der Mensch bringt ein Verständnis für den Kontext mit, das KI-Systemen schlicht fehlt.

Die Gefahr besteht, dass wir uns zunehmend auf KI verlassen, selbst in Bereichen, wo das kritisch ist. Wenn KI-Modelle in wichtige Entscheidungsprozesse integriert werden, müssen wir sicherstellen, dass sie von Menschen überwacht werden. Ein Beispiel: Ein KI-System könnte vorschlagen, eine bestimmte Brücke zu überqueren, obwohl alle anderen Menschen in die entgegengesetzte Richtung laufen. Ein Mensch würde diese Diskrepanz bemerken und hinterfragen, was die anderen Menschen sehen, was die KI möglicherweise übersehen hat. Solche Plausibilitätschecks sind unerlässlich, um Fehler zu vermeiden. Eine andere Gefahr ist, dass wir uns zu sehr auf KI-Systeme verlassen. So könnten sich Ärzte auf die KI verlassen und selbst nicht mehr genau auf die Tests und Laborwerte schauen. Das geschieht derzeit bei selbstfahrenden Autos. Der Fahrer ist verpflichtet, die Hände am Steuer zu haben, obwohl das Auto selbst fahren kann. Viele nehmen aber dann die Hände vom Steuer, da das Auto bisher alles korrekt gemacht hat. Das kann jedoch gefährlich sein, wenn eine unerwartete Situation eintritt.

Der Bildungssektor muss sich anpassen, um die nächste Generation auf eine Welt vorzubereiten, in der KI allgegenwärtig ist. Dabei geht es nicht nur darum, digitale Kompetenzen zu vermitteln, sondern auch darum, die einzigartigen Stärken des Menschen zu fördern – Kreativität, kritisches Denken und die Fähigkeit, Prinzipien zu abstrahieren. Ein Schwerpunkt sollte darauf liegen, Kindern und Erwachsenen beizubringen, wie sie KI als Werkzeug nutzen können, ohne von ihr abhängig zu werden.

Eine der größten Gefahren sehe ich in der politischen Manipulation durch KI. Systeme könnten gezielt genutzt werden, um Meinungen zu beeinflussen, Desinformation zu verbreiten oder gesellschaftliche Spaltungen zu vertiefen. Wenn diese Technologie in den Händen von unethischen Akteuren liegt, könnte sie erheblichen Schaden anrichten – etwa durch Fake News in sozialen Netzwerken, die Wähler beeinflussen sollen. Hier müssen wir als Gesellschaft klare Regeln und Kontrollmechanismen schaffen, um Missbrauch zu verhindern.

KI-Modelle können in sämtlichen Sprachen Nachrichten verbreiten und Stimmungen lenken, ohne dass ein Heer von Menschen an den Tastaturen sitzen muss. Hier tritt die KI gewissermaßen als »Schattenakteur« auf: Sie kann manipulieren – im schlimmsten Fall ganze Wahlkämpfe verzerren. Sie kann aber auch helfen, solche Falschmeldungen zu entlarven. »Gute« KI erkennt Deepfakes, Spam-Kampagnen oder koordinierte Troll-Angriffe und kann uns so vor gezielter Desinformation schützen. Hier besteht ein ständiges Wettrennen zwischen Angriffs- und Abwehrsystemen, zwischen »böser« und »guter« KI.

Empathische Maschinen?

Ein wiederkehrendes Thema in der Diskussion über künstliche Intelligenz ist die Frage, ob KI jemals menschliche Empathie haben kann. Neil Lawrence von der Universität Cambridge bringt dies in seinem Buch *The atomic human* treffend auf den Punkt: Maschinen fehlen menschliche

Schwachstellen – aber auch menschliche Stärken wie Empathie. Diese Unfähigkeit, echte Emotionen zu fühlen oder zu verstehen, wirft die Frage auf, ob KI wirklich in der Lage sein kann, Entscheidungen mit moralischem oder emotionalem Gewicht zu treffen.

Maschinen können keine Empfindungen wie Hunger, Schmerz, Trauer oder Freude erleben. Sie können nur Daten verarbeiten, die auf menschlichen Beschreibungen oder Verhaltensweisen basieren. Wenn ein Mensch zum Beispiel eine Strafe erhält, könnte eine KI die durchschnittliche Reaktion auf eine ähnliche Strafe analysieren, aber sie könnte niemals verstehen, wie sich diese Strafe für die betroffene Person tatsächlich anfühlt. Empathie ist für Menschen oft nicht nur ein rein rationaler Prozess, sondern basiert auf der Erfahrung ähnlicher Gefühle. Diese Erfahrungsebene fehlt der KI völlig.

Ein weiterer Aspekt ist, dass Gefühle und Reaktionen individuell und kulturell geprägt sind. Was für eine Person tragisch ist, könnte für eine andere unbedeutend sein. Diese Nuancen kann eine KI nur sehr schwer nachvollziehen, da dies nicht in den Daten vorhanden war, an denen die KI trainiert wurde.

Eine KI könnte theoretisch trainiert werden, bestimmte Reaktionen und Entscheidungen von Menschen nachzuahmen. Wenn beispielsweise Richter bei bestimmten Straftaten typischerweise milde Strafen verhängen, könnte eine KI dies analysieren und in ihren Entscheidungen berücksichtigen. Doch dabei handelt es sich um reine Mustererkennung, nicht um echtes Verstehen. Ein zentraler Punkt ist, dass KI niemals die zugrunde liegende Bedeutung dieser Muster verstehen wird. Sie könnte »lernen«, dass der Verlust eines geliebten Haustiers für manche Menschen eine immense emotionale Belastung darstellt, aber sie wird nicht nachvollziehen können, warum das so ist. Diese mangelnde Tiefe im Verständnis wird immer eine Grenze für KI darstellen.

Neil Lawrence schreibt in seinem Buch, dass die Art und Weise, wie Maschinen Geschichten erzählen oder menschliche Entscheidungen simulieren, fundamental anders ist als das, was uns Menschen ausmacht.

Hier liegt der Kern der Diskussion: Maschinen können Muster erkennen und auf dieser Basis Entscheidungen vorschlagen, aber sie können nicht die moralischen oder emotionalen Dimensionen dieser Entscheidungen wirklich begreifen.

Ein anschauliches Beispiel ist die Medizin. Ein Arzt muss häufig schwierige Entscheidungen treffen, etwa ob eine Behandlung sinnvoll ist, wenn sie die Schmerzen des Patienten verschlimmert, oder ob eine andere Behandlung mit weniger Schmerzen besser wäre. Während eine KI Daten analysieren und Wahrscheinlichkeiten berechnen kann, bleibt die moralische Abwägung – was für den Patienten das Beste ist – eine zutiefst menschliche Entscheidung.

Ein weiteres Beispiel sind Entscheidungen im Pflegebereich, etwa wenn es darum geht, ob eine ältere Person in ihrem Zuhause bleiben oder in ein Pflegeheim ziehen sollte. Hier spielen emotionale und soziale Faktoren eine große Rolle, die eine KI nicht erfassen kann. Nur ein Mensch kann wirklich einschätzen, was diese Veränderung für das Wohlbefinden der betroffenen Person bedeutet.

Auch wenn KI uns in vielen Bereichen unterstützen kann, gibt es eine klare Grenze: Entscheidungen, die tief in menschliche Emotionen und Werte eingebettet sind, sollten immer von Menschen getroffen werden. Eine KI kann Vorschläge machen oder Szenarien modellieren, aber sie sollte niemals die endgültige Entscheidung treffen. Dies ist besonders wichtig in Bereichen, die die Lebensqualität und das Wohlbefinden von Menschen direkt betreffen.

Ich finde es spannend, was Neil Lawrence in seinem Buch über die Assimilation von Menschlichkeit durch Maschinen sagt. Sein Gedanke, dass Maschinen Aspekte der Menschlichkeit »assimilieren«, also nicht nur imitieren, sondern sich an menschliche Muster anpassen, hat mich sehr beschäftigt. Für mich bedeutet das nicht, dass Maschinen wirklich menschlich werden, sondern dass sie sehr gut darin sind, menschliches Verhalten nachzuahmen. Sie wirken menschlich – das ist aber nicht dasselbe wie echte Menschlichkeit.

Die Large Language Models (LLMs), wie wir sie heute kennen, sind beeindruckend. Sie erwecken oft den Anschein, sie könnten »denken« oder sogar »fühlen«, aber das ist nur eine Illusion. Was diese Modelle tatsächlich tun, ist, menschliche Textdaten zu verarbeiten. Alles, was sie produzieren, basiert auf den menschlichen Sätzen, die sie in ihren Trainingsdaten hatten. Es gibt keine fremdartige Intelligenz, wie einige behaupten. Diese Modelle sind zutiefst menschenzentriert, weil sie vollständig auf menschlicher Sprache, menschlicher Logik und menschlichem Wissen basieren. Sie haben keine Erfahrung mit der physischen Welt.

KI-Systeme können viel über den Menschen lernen, schreibt Lawrence – aber sie lernen es ganz anders als wir. Das sehe ich genauso. Während wir durch Erfahrungen, Emotionen und Reflexion lernen, basieren Maschinen nur auf Algorithmen und Datenmustern. Sie können keine genuine Erfahrung machen. Für uns Menschen sind Geschichten nicht nur Worte – sie sind mit Gefühlen und Erinnerungen verknüpft. Eine Maschine hingegen kennt weder Freude noch Trauer, weder Stolz noch Verlust. Sie verarbeitet lediglich Muster, die sie in Daten gefunden hat. Gefühle wie Freude, Trauer, Stolz oder Verlust kommen aus unserer Evolution. Sie waren wichtig, um soziale Strukturen aufzubauen und in einer Gesellschaft leben und überleben zu können. Maschinen mussten nie in der Welt überleben und soziale Strukturen aufbauen.

Das bringt mich zu einem zentralen Punkt: Was Maschinen lernen, ist rein mechanisch. Sie können Geschichten neu erzählen, aber sie verstehen sie nicht wirklich. Das unterscheidet sie fundamental von uns Menschen. Der Prozess des Lernens und Verstehens ist bei uns untrennbar mit unseren Erfahrungen verbunden – mit all den Freuden, Leiden und Herausforderungen, die uns prägen. Maschinen fehlt dieser Bezug zur Realität völlig. Textgenerierungsmaschinen haben nie in der physischen Welt gelebt und können keinen Zusammenhang zur physischen Welt herstellen.

Eine oft gestellte Frage ist, was dem Menschen bleibt, wenn Maschinen in immer mehr Bereichen besser werden. Ich sehe das nicht als Bedrohung.

Vielmehr glaube ich, dass KI uns helfen kann, unsere einzigartigen menschlichen Fähigkeiten weiterzuentwickeln. Kreativität, Empathie und unsere Fähigkeit, moralische Entscheidungen zu treffen, bleiben Kernaspekte dessen, was uns ausmacht.

Was uns als Menschen wirklich einzigartig macht, ist unsere Fähigkeit, Verletzlichkeit zu erkennen und darauf zu reagieren. Das ist zentral für uns Menschen, die in Gruppen und in einer Gesellschaft leben – ein wichtiger sozialer Aspekt. Diese Eigenschaft wird keine Maschine je replizieren können. Deshalb sehe ich KI nicht als Konkurrenz, sondern als Ergänzung – ein Werkzeug, das uns hilft, die Welt besser zu verstehen und zu gestalten.

KI ist ein mächtiges Werkzeug, das bei richtiger Nutzung enorme Vorteile bringen kann. Doch wir dürfen uns nicht blenden lassen: KI hat Grenzen, insbesondere wenn es um Verständnis, Kontext und Kausalität geht. Der Mensch muss weiterhin die Kontrolle behalten, insbesondere in sensiblen Bereichen wie Politik, Finanzen, Medizin und Sicherheit. Denn nur wir verstehen soziale und gesellschaftliche Aspekte. Die Verantwortung liegt bei uns, KI sinnvoll einzusetzen und klare Leitlinien für ihre Nutzung zu schaffen. Nur so können wir das Potenzial der Technologie ausschöpfen, ohne ihre Gefahren zu übersehen.

Eines der größten Probleme im Umgang mit KI-Systemen ist das Spannungsfeld zwischen zu wenig und zu viel Vertrauen. Gerade in der Industrie habe ich erlebt, wie sich diese Extreme abspielen. Da gibt es auf der einen Seite die Skeptiker, die KI grundsätzlich misstrauen und sie trotz klarer Erfolge nicht ernst nehmen. Auf der anderen Seite stehen diejenigen, die sich blind auf KI verlassen, oft mit problematischen Konsequenzen.

Ein Beispiel aus der Pharmaindustrie: Ein Chemiker hatte festgestellt, dass KI-gestützte Vorhersagen tatsächlich zutrafen, als es darum ging, chemische Veränderungen vorzunehmen, um ein Medikament zu verbessern. Beeindruckt von den Ergebnissen, wiederholte er den Prozess immer wieder, ließ die KI das Design des Medikaments kontinuierlich optimieren. Doch irgendwann ging das Modell so weit über den ursprünglichen Trainingsdatensatz hinaus, dass die Ergebnisse unbrauchbar

wurden. Die KI begann, Strukturen vorzuschlagen, die in der Realität nicht mehr funktionierten, da solche Strukturen nie im Training waren. Hier zeigte sich die Gefahr des blinden Vertrauens: Man darf die Grenzen von KI nicht vergessen. Wenn Modelle auf Daten trainiert sind, können sie außerhalb dieser bekannten Bereiche keine verlässlichen Vorhersagen mehr treffen.

Dieses Phänomen, das ich in der Pharmaindustrie gesehen habe, lässt sich auch auf andere Bereiche übertragen – etwa auf selbstfahrende Autos. Zu Beginn ist man skeptisch und bleibt wachsam, hält die Hände am Lenkrad und greift ein, wenn nötig. Doch mit der Zeit, wenn alles reibungslos funktioniert, wird man nachlässig. Man vertraut dem System immer mehr, bis man am Ende vielleicht sogar die Hände vom Steuer nimmt. Doch was passiert, wenn das System eine unerwartete Situation falsch einschätzt? Dieses »automatische Vertrauen« ist ein Problem, das weit über den Automobilsektor hinausgeht.

Auch in anderen Branchen sehe ich oft, dass Menschen dazu neigen, KI-Systeme entweder komplett abzulehnen oder ihnen mehr zuzutrauen, als sie tatsächlich können. Beide Extreme sind problematisch. Es ist entscheidend, dass wir lernen, KI als Werkzeug zu sehen – eines, das uns unterstützt, das uns hilft, die Welt besser zu verstehen und zu gestalten, das uns aber nicht ersetzt!

11

Warum wir KI brauchen

Wie können wir Krebs heilen? Wie können wir den Klimawandel aufhalten? Wie können wir sicherstellen, dass alle Menschen Zugang zu Bildung, Nahrung und Gesundheitsversorgung haben?

Das sind die Fragen, auf die wir uns konzentrieren sollten – und auf die Bereiche, in denen KI uns wirklich helfen kann, Lösungen für die großen Probleme der Menschheit zu finden.

Die Frage, welche Rolle KI in unserer Zukunft spielen sollte, ist für mich zentral. Ich denke, KI ist nicht nur ein »Assistent« oder ein »Helferlein«, der uns nur Bequemlichkeit bringt, indem er E-Mails schreibt oder kleinere Aufgaben erledigt. Nein, KI hat das Potenzial, die wirklich großen Probleme der Menschheit anzugehen.

Die Herausforderungen, vor denen wir stehen, beinhalten zu komplexe Systeme, als dass ein Mensch sie allein durchdenken könnte. Denken wir an die Wechselwirkungen beim Klimawandel – Hunderte Faktoren, die sich gegenseitig beeinflussen. Oder an die Entwicklung neuer Medikamente: Die Wechselwirkungen in und zwischen menschlichen Zellen sind so komplex, dass sie selbst mit herkömmlichen Computermodellen nicht zu erfassen sind.

Hier sehe ich die Stärke der KI. Sie kann uns helfen, diese hochkomplexen Systeme zu analysieren, Szenarien zu modellieren und Lösungen zu finden, die unsere intellektuellen Fähigkeiten weit übersteigen.

Auch wirtschaftliche und geopolitische Zusammenhänge werden zunehmend komplexer. Politische Entscheidungen in einem Land haben Auswirkungen auf die Weltwirtschaft, auf Lieferketten und Energiepolitik. KI kann helfen, diese Dynamiken zu verstehen und fundierte

Entscheidungen zu treffen. Dennoch bleibt es wichtig, dass der Mensch die Verantwortung behält.

Themen wie die weltweiten Gesundheitssysteme, die Bekämpfung von Krankheiten und Pandemien, unsere zunehmend vernetzte Weltwirtschaft und natürlich der Klimawandel stehen ganz oben auf der Liste. Ich bin fest davon überzeugt, dass KI in diesen Bereichen enorm helfen kann – sei es bei der Entwicklung neuer Materialien, in der Energieoptimierung oder im intelligenten Ressourcenmanagement.

Gerade im Klimabereich setze ich große Hoffnungen in die KI. Durch smarte Steuerungssysteme im Haushalt lässt sich beispielsweise viel Energie sparen: Das Licht kann automatisiert ausgeschaltet werden, wenn keiner daheim ist, oder die Heizung wird sofort heruntergeregelt, sobald die Bewohner das Haus verlassen. Selbst in Hotels, wo immer noch massiv Energie verschwendet wird, könnten KI-basierte Modelle lernen, wann es sich überhaupt lohnt, Räume zu beheizen oder zu beleuchten.

Ein weiterer Aspekt ist unsere verwobene Weltwirtschaft. Wir sehen bei politischen Konflikten oder Kriegen, wie stark die Produktion und Lieferung bestimmter Güter plötzlich einbrechen kann. Auch Themen wie die Energiepolitik sind komplex: Sobald ein Staat ein Atomkraftwerk hochfährt, ändert sich der Strommix, die Nachbarländer reagieren, kaufen günstigen Strom – oder dürfen es nicht, weil sie sich selbst verpflichtet haben, einen Mindestanteil an »grünem Strom« zu haben. Dieses globale Miteinander ist derart verknüpft, dass ohne KI kaum mehr zu durchschauen ist, wer wann und wo Einfluss auf wen nimmt.

Aus meiner Sicht ist KI deshalb mehr als nur »nice to have«: Sie wird zum entscheidenden Werkzeug, um solche dynamischen Wechselwirkungen und Vernetzungen zu erfassen und bessere Entscheidungen zu treffen. Wir dürfen dabei allerdings nicht vergessen, dass die KI zwar beste Ergebnisse liefern kann, aber die Entscheidung, wie wir damit umgehen, nach wie vor bei uns Menschen liegt.

Die Rolle der KI in der Wissenschaft wird immens sein, und zwar über alle Disziplinen hinweg. Oft wird nun von »AI in Science« gesprochen und

viele Projekte werden in diese Richtung aufgesetzt. KI wird nicht nur Simulationen und Modellierungen verbessern, sondern auch als Werkzeug dienen, um wissenschaftliche Prozesse effizienter zu gestalten. Von der Suche nach Literatur über die Generierung von Hypothesen bis hin zur Automatisierung von Experimenten – KI wird die Art und Weise, wie wir Wissenschaft betreiben, grundlegend verändern.

Der Mensch wird dabei immer eine zentrale Rolle spielen, weil er die Richtung vorgeben muss. Es braucht uns, um die wichtigen Fragen zu stellen: Welche Krankheiten sollen bekämpft werden? Welche physikalischen Phänomene wollen wir besser verstehen? Auch wenn die KI viele Details übernimmt, bleibt die kreative, richtungsweisende Arbeit beim Menschen.

Eine spannende Frage ist, ob KI in der Lage sein wird, echte wissenschaftliche Entdeckungen zu machen. Ich glaube, sie wird uns dabei helfen, Entdeckungen zu machen, aber der Mensch wird immer derjenige sein, der die entscheidenden Impulse gibt. Selbst bei einer so komplexen Fragestellung wie der »Riemannschen Vermutung«, einem der größten ungelösten Probleme der Mathematik, könnte eine KI möglicherweise den Beweis liefern. Aber die Entscheidung, diese Frage überhaupt anzugehen, muss von einem Menschen kommen.

Die Zukunft der Wissenschaft mit KI ist aufregend. KI wird uns ermöglichen, komplexe Systeme zu modellieren, Hypothesen zu testen und neue Entdeckungen zu machen. Doch KI wird uns nicht ersetzen – sie wird uns ergänzen. Der Mensch bleibt der kreative Motor, der die Richtung vorgibt und die entscheidenden Fragen stellt.

Am Ende ist es vielleicht genau das: KI hilft uns, unsere eigenen intellektuellen Grenzen zu überschreiten. Sie wird nicht alles lösen können, aber sie wird uns zeigen, wo wir selbst nicht weiterkommen – und uns so ermöglichen, die Welt noch besser zu verstehen.

Als Menschen sind wir evolutionär darauf ausgelegt, in kleinen Gruppen zu überleben, einfache soziale Interaktionen zu managen und grundlegende Probleme zu lösen. Doch die Welt, in der wir heute leben, ist weitaus

komplexer. In Bereichen wie der Genetik oder der Klimawissenschaft stoßen wir an Grenzen. Es gibt Phänomene, die wir zwar beobachten, aber nicht vollständig verstehen können, weil sie aus unzähligen Interaktionen bestehen. Hier könnte KI uns helfen, Muster und Zusammenhänge zu entdecken, die für uns unerkannt bleiben.

Vielleicht ist die wahre Stärke der KI nicht, dass sie »intelligenter« ist als wir, sondern dass sie völlig anders an Probleme herangeht. Sie könnte uns neue Wege zeigen, über Probleme nachzudenken, die wir niemals allein finden würden. Diese »andere Intelligenz« könnte uns helfen, die Welt auf eine Weise zu verstehen, die bisher jenseits unserer Vorstellungskraft liegt.

Die Diskussion, ob uns Menschen bei vielen Phänomenen das Konzept fehlt oder ob wir die Dinge schlicht nicht überblicken können, finde ich faszinierend. Ich denke, es ist oft eine Kombination aus beidem. In manchen Fällen ist ein System so komplex, dass wir keinen Zugang zu einer vereinfachten Darstellung finden können – eben weil uns der Überblick fehlt. Und ohne diesen Überblick können wir auch kein passendes Konzept entwickeln.

Was ich nicht will, ist eine KI, die lediglich die Bequemlichkeit fördert. Ich nenne das immer »mechanische Sklaverei« – eine KI, die unsere alltäglichen Aufgaben übernimmt, uns noch fauler macht und uns letztlich von unseren eigenen Fähigkeiten entfremdet. Wir brauchen eine KI, die uns über unsere eigenen Grenzen hinausführt, nicht eine, die uns in unseren Komfortzonen verharren lässt. Dabei ist mir auch die Frage wichtig, was passiert, wenn wir zu bequem werden. Die Vorstellung, dass KI uns alles abnimmt, mag verlockend sein, aber ich sehe darin auch eine Gefahr. Es gibt dystopische Szenarien, in denen Menschen durch ihre eigene Bequemlichkeit zugrunde gehen. Wenn wir uns nicht mehr bewegen, nicht mehr selbst denken und keine Herausforderungen mehr haben, verlieren wir einen wesentlichen Teil dessen, was uns als Menschen ausmacht. Ein Leben, in dem alles perfekt organisiert und vorhergesagt ist, klingt auf den ersten Blick attraktiv, kann aber schnell zu einer Dekadenz führen, die uns als Gesellschaft schwächt. Insbesondere denke ich hier auch an soziale

Interaktionen zwischen Menschen, die durch KI eventuell verringert werden könnten.

Ich glaube, dass Herausforderungen und Anstrengungen uns als Menschen weiterbringen. Bewegung, geistige Anstrengung und soziale Interaktion sind keine Last, sondern etwas, das unser Leben bereichert. Wenn wir diesen Aspekt aus unserem Leben entfernen, verlieren wir einen Teil unserer Menschlichkeit. KI sollte uns dabei unterstützen, ein besseres Leben zu führen – nicht ein bequemeres auf Kosten unserer Vitalität und Kreativität.

Die Versuchung, alles automatisieren zu lassen, ist groß. Die KI liest für uns, schreibt für uns, entscheidet für uns. Aber wohin führt das? Wir verlieren nach und nach unsere Fähigkeiten, weil wir sie nicht mehr nutzen. Schreiben, Lesen, selbst Nachdenken – all das sind Tätigkeiten, die uns nicht nur als Individuen, sondern auch als Gesellschaft formen. Wenn wir diese Tätigkeiten an Maschinen delegieren, geben wir nicht nur Aufgaben ab, sondern auch einen Teil unserer Identität.

Es gibt Tätigkeiten, die auf den ersten Blick mühsam wirken – aber sie bereichern uns. Ein Buch selbst zu lesen, anstatt sich die Zusammenfassung vorlesen zu lassen, bedeutet, die Feinheiten und Nuancen der Sprache wahrzunehmen. Kochen bedeutet nicht nur Nahrung zubereiten, sondern auch Kreativität und Stolz auf das Geschaffene. Diese Erfahrungen machen uns zu dem, was wir sind. Sie sollten nicht der Bequemlichkeit geopfert werden.

Auch in der Wissenschaft sollte man agil und frisch bleiben. Ich sehe es als meine Verantwortung, jungen Forschern zu sagen: »Lasst euch nicht entmutigen!« Es mag manchmal so aussehen, als ob alles schon erforscht ist, oder als ob nur die großen Firmen mit ihren riesigen Budgets Innovationen schaffen können. Aber das stimmt nicht. Gute Ideen sind unbezahlbar. Eine brillante Idee kann mehr wert sein als Milliarden Euro. Das habe ich selbst anhand von LSTM erlebt. Dieses Gefühl, etwas geschaffen zu haben, das funktioniert, ist unvergleichlich. Es hält länger an und ist erfüllender als Ruhm, Macht oder Reichtum. Genau das möchte ich

den nächsten Generationen weitergeben. Man sieht das nun wieder bei DeepSeek: Gute Ideen zusammen mit der richtigen Infrastruktur können etwas bewegen.

Für mich liegt der Schlüssel in der Balance. KI sollte uns bei den Dingen unterstützen, die uns belasten, und uns Freiräume schaffen, um uns auf das Wesentliche zu konzentrieren. Aber sie sollte uns nicht unserer Eigenständigkeit berauben. Ich sehe KI als ein Tool, das uns von zeitraubenden und monotonen Aufgaben befreit. Steuererklärungen, Buchhaltung, einfache Planungen – hier kann sie uns Arbeit abnehmen, ohne dass wir etwas Wesentliches verlieren. Aber bei anderen Dingen, wie dem Lesen eines Buches, dem Kochen oder dem Schachspielen, ist es anders. Diese Tätigkeiten bereichern uns – sie machen uns zu dem, was wir sind.

Wenn KI uns die langweiligen und uninteressanten Aufgaben abnimmt, gewinnen wir Zeit für die wirklich wichtigen Dinge im Leben: für Familie und Freunde, für unsere Gesundheit, für Kreativität und Reflexion. Das ist der wahre Wert der Technologie.

Ich möchte, dass KI uns unterstützt, aber nicht ersetzt. Sie soll uns helfen, Krankheiten zu heilen, die Umwelt zu schützen und die großen Probleme der Welt zu lösen. Aber sie soll uns auch Raum lassen, um zu wachsen, zu lernen und die Welt selbst zu entdecken.

Eine besonders faszinierende Vision ist die eines »universellen Wissensmodells«. Ein solches Modell würde das gesamte Wissen der Menschheit integrieren – von physikalischen Gesetzmäßigkeiten über historische Ereignisse bis hin zu aktuellen wissenschaftlichen Erkenntnissen. Mit einem solchen Modell könnten Menschen auf jede erdenkliche Frage eine fundierte Antwort erhalten. Es würde den Zugang zu Wissen demokratisieren und die Art und Weise, wie wir lernen und forschen, grundlegend verändern.

Doch dieses Ziel ist noch weit entfernt. Es erfordert nicht nur immense Datenmengen, sondern auch die Fähigkeit der KI, diese Daten sinnvoll zu strukturieren und zu interpretieren. Darüber hinaus muss ein solches Modell

in der Lage sein, zwischen Fakten und Fiktion zu unterscheiden – eine Herausforderung, die angesichts der Fülle an Fehlinformationen im Internet sehr groß ist. Die jetzigen LLMs gehen in die Richtung, das Wissen der Menschheit zu speichern, aber solange sie halluzinieren, unvollständiges Wissen haben und nicht logisch schließen können, sind wir nicht am Ziel. Außerdem ist das Wissen der Menschen nicht nur in Textform vorhanden, da viel in Datenbanken oder Verzeichnissen abgelegt ist.

Wenn ich an die nächste Generation denke, wünsche ich mir eine Welt, in der sich Menschen für etwas begeistern können. Es muss nicht Mathematik oder KI sein. Es kann Musik, Sport oder etwas völlig anderes sein. Wichtig ist, dass sie eine Leidenschaft finden, die sie antreibt. Und ich hoffe, dass sie in einer Welt leben, in der KI sie unterstützt, ohne sie zu erdrücken.

Meine Vision ist eine Zukunft, in der Technologie nicht nur unsere Arbeit erleichtert, sondern auch unsere Beziehungen bereichert. Eine Welt, in der wir wieder mehr Zeit füreinander haben, weil uns die Maschinen den Rücken freihalten. Und eine Welt, in der junge Menschen erkennen, dass ihre Ideen wertvoller sind als jeder technologische Fortschritt. Das ist für mich der Kern dessen, was uns als Menschen ausmacht: unsere Fähigkeit zu träumen, zu gestalten und miteinander verbunden zu bleiben.

GLOSSAR

Abstraktion

Der Prozess, komplexe Informationen auf wesentliche Merkmale zu reduzieren, um sie effizient zu verarbeiten und auf verschiedene Situationen und Kontexte anwendbar zu machen. Weglassen von Einzelheiten und das Überführen auf etwas Allgemeineres oder Einfacheres.

AGI (Artificial General Intelligence)

Hypothetische Form künstlicher Intelligenz, die kognitive Fähigkeiten besitzt, so wie wir Menschen sie haben.

Aktivierung

Die Berechnung des Ausgabewerts eines Neurons in einem neuronalen Netz mithilfe einer Aktivierungsfunktion, die entscheidet, ob und in welchem Maß ein Signal weitergegeben wird.

Attention-Mechanismus

Dieser Mechanismus ermöglicht es neuronalen Netzen, sich gezielt auf die relevantesten Teile einer Eingabesequenz zu konzentrieren, anstatt alle Elemente der Sequenz gleich zu gewichten. So erfasst der Self-Attention-Mechanismus in Transformern den paarweisen Zusammenhang zwischen weit entfernten Wörtern in einem Text.

Backpropagation

Ein effizientes Berechnungsverfahren für den Gradienten von Fehlern nach den Gewichten bei neuronalen Netzen. Der Gradient wird dann für das Gradientenabstiegsverfahren verwendet, um das Netz zu optimieren. Dabei werden die Gewichte des Netzes schrittweise angepasst, um es zu verbessern. Der Fehler einer Vorhersage wird durch Anpassung der Gewichte minimiert.

Chain of Thought (CoT)

Eine Technik zur Verbesserung der Entscheidungsfindung in großen Textmodellen, bei der das Modell explizit Zwischenschritte in der Argumentation ausgibt. Es teilt komplexe Probleme in einfachere Teilprobleme auf, ähnlich wie Menschen in Teilschritten denken.

Convolutional Neuronal Networks (CNN)

Neuronale Netze, die besonders gut für die Verarbeitung von Bild- und Videodaten geeignet sind.

Deep Learning

Ein Teilbereich des maschinellen Lernens, der künstliche neuronale Netze mit vielen Schichten nutzt. Es erkennt komplexe Muster in großen Datenmengen und löst damit schwierige Aufgaben. Es bildet die Grundlage der modernen KI und wird in Bild-, Text- und Sprachverarbeitung eingesetzt.

Digitaler Zwilling

Ein virtuelles Simulationsmodell eines physischen Prozesses, das in Echtzeit mit seinem realen Gegenstück gekoppelt ist. Durch die Simulation mit dem digitalen Zwilling können Prozesse in Echtzeit gesteuert und Entscheidungsprozesse in Echtzeit verbessert werden.

Expertensysteme

KI-Systeme, die auf regelbasiertem Wissen und logischen Schlussfolgerungen basieren, um menschliche Experten in spezifischen Bereichen zu unterstützen.

Foundation Models

Große, vortrainierte KI-Modelle, die auf enormen Datenmengen basieren und für vielfältige Anwendungen angepasst werden können. Sie bilden die Grundlage für spezialisierte KI-Anwendungen in Textverarbeitung, Bildverarbeitung, Programmierung und anderen Bereichen.

Garantierbarkeit

Die Fähigkeit eines KI-Systems, zuverlässige und überprüfbare Ergebnisse mit vorhersehbarer Genauigkeit zu liefern, besonders relevant für sicherheitskritische Anwendungen.

Gate

Eine grundlegende Recheneinheit in neuronalen Netzen, die bestimmt, welche Informationen weitergegeben oder blockiert werden, z. B. in LSTMs zur Steuerung des Informationsflusses.

Generalisierung

Die Fähigkeit eines KI-Modells, Muster aus Trainingsdaten auf neue, unbekannte Daten zu übertragen, ohne dabei spezifische Details der Trainingsdaten auswendig zu lernen.

Gewichte (Weights)

Parameter eines neuronalen Netzes, die die Stärke der Verbindung zwischen Neuronen bestimmen. Während des Trainings werden die Gewichte durch Backpropagation angepasst, um den Trainingsfehler des Modells zu verkleinern.

GPT (Generative Pre-trained Transformer)

Ein KI-Modell basierend auf der Transformer-Architektur, das Texte generiert, indem es Informationen über zuvor trainierte Texte in seinen Gewichten speichert. Es wird in zwei Phasen trainiert: allgemeines Pre-Training und spezifische Anpassung (Fine-Tuning).

GPU (Graphics Processing Unit)

Ein spezialisierter Prozessor für parallele Matrixmultiplikation, der für parallele Berechnungen optimiert ist und daher besonders für KI-Modelle und große Datenmengen eingesetzt wird. GPUs beschleunigen das Training neuronaler Netze erheblich im Vergleich zu herkömmlichen CPUs.

Gradient

Mathematische Ableitung, die angibt, wie stark eine Änderung der Gewichte in einem neuronalen Netz den Fehler beeinflusst; entscheidend für das Training mittels Gradientenabstiegs.

Hidden Spaces oder Embedding Spaces

Versteckte Schichten (Hidden Layers) oder Einbettungsräume in einem neuronalen Netz, die abstrakte Repräsentationen von Daten durch komplexe Merkmale enthalten.

Hidden Layers (Versteckte Schichten)

Zwischenschichten in neuronalen Netzen, die zwischen Eingabe- und Ausgabeschicht liegen und komplexe Merkmale aus der Eingabe extrahieren, essenziell für Deep Learning.

Hopfield-Netzwerke

Ein sehr alter spezieller Typ rekurrenter neuronaler Netze, der als assoziativer Speicher dient und Muster auch bei unvollständigen

Eingaben rekonstruieren kann.

Inferenz

Der Prozess, bei dem ein trainiertes KI-Modell eingesetzt wird. Das Modell trifft neue Vorhersagen oder Schlussfolgerungen auf Basis gegebener Eingaben. In neuronalen Netzen bedeutet dies, dass das Modell auf Basis der trainierten Gewichte aus einer Eingabe eine Ausgabe generiert.

Large Language Models (LLMs)

Große, meist transformerbasierte KI-Textmodelle, die durch Training auf umfangreichen Textdaten Texte generieren können, die von menschlichen Texten kaum zu unterscheiden sind (z. B. GPT-4).

Long Short-Term Memory (LSTM)

Ein spezieller Typ rekurrenter neuronaler Netze (RNNs), der durch Gating-Mechanismen und einen speziellen Speicher lernen kann, Informationen über eine längere Zeit zu speichern.

Matrix

Rechteckige Anordnungen von Zahlen, die in der KI und im maschinellen Lernen zur Darstellung und Berechnung von Gewichten in neuronalen Netzwerken genutzt werden.

Matrizenmultiplikation

Eine grundlegende Rechenoperation in der KI, bei der eine Matrix mit einem Vektor kombiniert wird, um Transformationen durchzuführen, insbesondere zur Berechnung von Aktivierungen und Gradienten in neuronalen Netzen.

Monte Carlo Tree Search (MCTS)

Ein probabilistischer Suchalgorithmus für Entscheidungsbäume, der durch einen zufälligen Abstieg im Baum die besten möglichen Züge ermittelt. Er wird häufig in Spielen wie Schach oder Go eingesetzt.

Neuronale Netze

KI-Modelle, die aus künstlichen Neuronen bestehen, die über gewichtete Verbindungen Signale austauschen. Ausgehend von einer Eingabe aktivieren Neuronen durch die Verbindungen andere Neuronen und generieren so komplexe Muster und Zusammenhänge aus den Eingabedaten. Mit diesen Mustern löst das neuronale Netz eine Aufgabenstellung, etwa Hundebilder von Katzenbildern zu unterscheiden.

Neuronen

Grundbausteine neuronaler Netze, die über eingehende Verbindungen von anderen Neuronen aktiviert werden und über ausgehende Verbindungen Signale an andere Neuronen weiterleiten. Die Aktivierungen der Neuronen repräsentieren Muster in der Eingabe, die gelernt wurden.

Overfitting

Ein Problem im maschinellen Lernen, bei dem ein Modell zu stark an Trainingsdaten angepasst wird und deshalb auf neuen Daten schlecht generalisiert. Es lernt etwa das Rauschen in den Trainingsdaten, das aber zufällig ist.

Open Source

KI-Systeme, deren Quellcode, Trainingsdaten und Modellparameter öffentlich zugänglich sind und von jedem genutzt, inspiziert, geändert und weiterverbreitet werden können.

Perzeptron

Das einfachste künstliche neuronale Netz mit einer Eingabe- und einer Ausgabeschicht mit nur einem Neuron, ursprünglich für binäre Klassifikationsaufgaben entwickelt.

Rechenkomplexität

Die theoretische Bewertung des Rechenaufwands eines Algorithmus in Bezug auf Zeit- und Speicherbedarf, entscheidend für effiziente KI-Modelle.

Reinforcement Learning (Bestärkendes Lernen)

Ein Lernverfahren, bei dem ein Agent durch Belohnung oder Bestrafung eigenständig optimale Handlungsstrategien erlernt.

Rekurrente Netze (RNNs)

Neuronale Netze, die Rückkopplungen enthalten, sodass frühere Eingaben in die Berechnung neuer Zustände einfließen und somit ein Gedächtnis haben. Sie sind besonders für Sequenzverarbeitung geeignet.

Simulation

Die Nachbildung realer physikalischer Prozesse über ein Zeitintervall in einer virtuellen Umgebung, um Vorhersagen zu treffen.

Skalierung

Die Fähigkeit eines KI-Systems, mit zunehmender Modellgröße, Datenmenge und Rechenleistung leistungsfähiger zu werden, entscheidend für den Erfolg von LLMs.

Subsymbolische KI

Ein KI-Ansatz, der Wissen nicht explizit durch Regeln, sondern durch Lernen anhand von Daten generiert. Er ist die Grundlage des maschinellen Lernens, insbesondere von neuronalen Netzen, d. h. von Deep Learning.

Support Vector Machines (SVMs)

Ein klassisches subsymbolisches Machine-Learning-Verfahren zur Klassifikation und Regression, das Entscheidungsgrenzen optimal trennt.

Symbolische KI

Ein Ansatz der künstlichen Intelligenz, der Wissen durch logische Regeln und Symbole darstellt, um Wissen zu repräsentieren und Schlussfolgerungen zu ziehen.

Transformer

Eine KI-Architektur, die durch Self-Attention-Mechanismen parallele Berechnungen ermöglicht und die Grundlage für LLMs bildet.

Vektor

Mathematische Größe mit Richtung und Betrag, die in KI zur Darstellung von Datenpunkten, Texten oder Bildern in mehrdimensionalen Räumen genutzt werden.

Quellenhinweise

Allen Institute for AI, Open Source LLM neben Llama von META

Team OLMo, P. Walsh, L. Soldaini, D. Groeneveld, K. Lo., S. Arora, A. Bhagia, Y. Gu, S. Huang, M. Jordan, N. Lambert, D. Schwenk, O. Tafjord, T. Anderson, D. Atkinson, F. Brahman, C. Clark, P. Dasigi, N. Dziri, M. Guerquin, H. Ivison., P. W. Koh, J. Liu, S. Malik, W. Merrill, L. J. V. Miranda, J. Morrison, T. Murray, C. Nam, V. Pyatkin, A. Rangapur, M. Schmitz, S. Skjonsberg, D. Wadden, C. Wilhelm, M. Wilson, L. Zettlemoyer, A. Farhadi, N. A. Smith, H. Hajishirzi: *2 OLMo 2 furious*, 2025. URL https://arxiv.org/abs/2501.00656.

ChatGPT Artikel

J. Schulman, B. Zoph, C. Kim, J. Hilton, et al.: *ChatGPT: Optimizing language models for dialogue.* https://openai.com/blog/chatgpt/, 2022. OpenAI Research.

Deep Learning Buch

I. Goodfellow, Y. Bengio, A. Courville: *Deep learning*. MIT Press. 2016.

Deep Learning Nature-Artikel

Y. LeCun, Y. Bengio, G. Hinton: *Deep learning*. Nature, 521(7553):436, 2015.

DeepSeek

DeepSeek-AI, D. Guo, D. Yang, H. Zhang, J. Song, R. Zhang, R. Xu, Q. Zhu, S. Ma, P. Wang, X. Bi, and et al.: *DeepSeek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning*, January 2025. URL http://arxiv.org/abs/2501.12948.ArXiv:2501.12948.

Einführung von Transformern 2017

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin: »Attention is all you need«. In: *Advances in Neural Information Processing Systems (NeurIPS)*, volume 30, pp. 5998–6008. Curran Associates, Inc., 2017.

Flashattention

T. Dao: *FlashAttention-2: Faster attention with better parallelism and work partitioning. In Proceedings of the International Conference on Learning Representations (ICLR)*, 2024. URL https://openreview.net/forum?id=mZn2Xyh9Ec. T. Dao, D. Y. Fu, S. Ermon, A. Rudra, and C. Ré: »Flashattention: Fast and memory-efficient exact attention with IO-awareness«. In: *A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho (eds.), Advances in Neural Information Processing Systems (NeurIPS)*, 2022. URL https://openreview.net/forum?id=H4DqfPSibmx.

Forget Gate

F. A. Gers, J. Schmidhuber, and F. Cummins: *Learning to forget: Continual prediction with LSTM. Neural Computation*, 12(10):2451–2471, 2000.

Google LLM Gemini

Gemini Team Google: *Gemini: A family of highly capable multimodal models.* ArXiv, 2312.11805, 2023.

Google's LLM Gemma

Gemma Team: *Gemma 2: Improving open language models at a practical size, 2024b*. URL https://arxiv.org/abs/2408.00118.

Gemma Team: *Gemma: Open models based on gemini research and technology.* 2024. URL https://arxiv.org/abs/2403.08295.

GPT-Modell und die neuen Fähigkeiten

A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever: *Language models are unsupervised multitask learners.* https://openai.com/index/betterlanguage-models, 2019.

Großes LLM von Google (war damals groß)

L. Xue, N. Constant, A. Roberts, M. Kale, R. Al-Rfou, A. Siddhant, A. Barua, and C. Raffel: »mT5: A massively multilingual pre-trained text-to-text transformer«. In: *Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 483–498, 2021. doi: 10.18653/v1/2021.naacl-main.41.

Hybrid-Architektur, hier Transformer mit Mamba: State-Space_Modell

M. Poli, A. W. Thomas, E. Nguyen, P. Ponnusamy, B. Deiseroth, K. Kersting, T. Suzuki, B. Hie, S. Ermon, C. Ré, C. Zhang, and S. Massaroli: *Mechanistic design and scaling of hybrid architectures*. ArXiv, 2403.17844, 2024.

Ilya Sutskever (OpenAI-Gründer) wurde bekannt mit LSTM-Übersetzung

I. Sutskever, O. Vinyals, and Q. V. V. Le: »Sequence to sequence learning with neural networks«. In: Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (eds.), *Advances in Neural Information Processing Systems 27 (NIPS'13)*, pp. 3104–3112. Curran Associates, Inc., 2014.

Learning to Learn: LSTM lernt einen Lernalgorithmus wie Backpropagation

S. Hochreiter, A. Steven Younger, and Peter R. Conwell: »Learning to learn using gradient descent«. In: G. Dorffner, H. Bischof, and K. Hornik (eds.), *Proc. Int. Conf. on Artificial Neural Networks (ICANN 2001)*, pp. 87–94. Springer, 2001.

Llama LLM von META

H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, A. Rodriguez, A. Joulin, E. Grave, and G. Lample. Llama: *Open and efficient foundation language models*. ArXiv, 2302.1397, 2023.

LLM von META mit 175 Milliarden Parametern, vergleichbar zu GPT-3 von OpenAI

S. Zhang, S. Roller, N. Goyal, M. Artetxe, M. Chen, S. Chen, C. Dewan, M. Diab, X. Li, X. V. Lin, T. Mihaylov, M. Ott, S. Shleifer, K. Shuster, D. Simig, P. S. Koura, A. Sridhar, T. Wang, and L. Zettlemoyer. OPT: *Open pretrained transformer language models.* ArXiv, 2205.01068, 2022.

LSTM

S. Hochreiter: *Untersuchungen zu dynamischen neuronalen Netzen.* Master's thesis, Technische Universität München, 1991.

S. Hochreiter and J. Schmidhuber: *Long short-term memory. Neural Computation*, 9(8):1735–1780, 1997.

LSTM Lernparameter und Architektursuche

K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber: *LSTM: A search space odyssey.* Ar-Xiv, 1503.04069, 2015.

LSTM Sequenzgenerierung

A. Graves: *Generating sequences with recurrent neural networks.* ArXiv, 1308.0850, 2013.

LSTM-Netz gewinnt Starcraft II

O. Vinyals, T. Ewalds, S. Bartunov, et al.: *Starcraft II: A new challenge for reinforcement learning*. ArXiv, 1708.04782, 2017.

Mamba als Alternative zu Transformern: State-Space-Modelle

A. Gu and T. Dao: *Mamba: Linear-time sequence modeling with selective state spaces. In International Conference on Learning Representations (ICLR)*, 2024. URL https://openreview.net/forum?id=AL1fq05o7H.

Mistral LLM

Mistral AI Team: *Codestral Mamba.* https://mistral.ai/news/codestral-mamba/, 2024.

NVIDIA LLM

M. Shoeybi, M. Patwary, R. Puri, P. LeGresley, J. Casper, and B. Catanzaro: *Megatron-LM: Training multi-billion parameter language models using model parallelism.* ArXiv, 1909.08053, 2019.

OpenAI Five: LSTM-Netzwerk hat Dota-Weltmeister geschlagen

A. Karpathy: *OpenAI Five defeats Dota 2 world champions.* https://openai.com/research/openai-five-defeats-dota-2world-champions, 2019.

OpenAI-Gründer wurde mit LSTM berühmt

A. Karpathy: *The unreasonable effectiveness of recurrent neural networks.* http://karpathy.github.io/2015/05/21/rnn-effectiveness/, 2015.

Paper, wo Wojciech Zaremba (OpenAI-Gründer) mit LSTM bekannt wurde

W. Zaremba and I. Sutskever: *Learning to execute.* ArXiv, 1410.4615, 2014.

ResNet-Architektur, die große Modelle möglich macht

K. He, X. Zhang, S. Ren, and J. Sun: »Deep residual learning for image recognition«. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016.

Scaling Laws: Größere Modelle werden besser

J. Kaplan, S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, S. Gray, A. Radford, J. Wu, and D. Amodei: *Scaling laws for neural language models.* ArXiv, 2001.08361, 2020.

Überblicksartikel von Schmidhuber zu Deep Learning: Zitieren für Deep Learning

J. Schmidhuber: *Deep learning in neural networks: An overview.* Neural Networks, 61:85–117, 2015. doi: 10.1016/j.neunet.2014.09.003.

Unterschriften von Bildern generieren mit LSTM

M. D. Hossain, F. Sohel, M. F. Shiratuddin, and H. Laga: *A comprehensive survey of deep learning for image captioning.* ACM Computing Surveys (CSUR), 51(6):118, 2019.

Vanishing Gradient mit Bengio schon in meiner Diplomarbeit

S. Hochreiter, Y. Bengio, P. Frasconi, and J. Schmidhuber: »Gradient flow in recurrent nets: The difficulty of learning long-term dependencies«. In: J. Kolen and S. Kremer (eds.), *A Field Guide to Dynamical Recurrent Networks.* IEEE, 2000.

xLSTM für Bildverarbeitung (ViL): Vision-xLSTM

B. Alkin, M. Beck, K. Pöppel, S. Hochreiter, J. Brandstetter: *»Vision-LSTM: xLSTM* as generic vision backbone«. In: *Proceedings of the International Conference on Learning Representations (ICLR)*, 2025. URL https://openreview.net/forum?id=SiH7DwNKZZ.

xLSTM für biologische Anwendungen

N. Schmidinger, L. Schneckenreiter, P. Seidl, J. Schimunek, P.-J. Hoedt, J. Brandstetter, A. Mayr, S. Luukkonen, S. Hochreiter, G. Klambauer: »Bio-xLSTM: Generative modeling, representation and in-context learning of biological and chemical sequences«. In: *Proceedings of the International Conference on Learning Representations (ICLR)*, 2025. URL https://openreview.net/forum?id=IjbXZdugdj.

xLSTM für Robotik und Agenten

T. Schmied, T. Adler, V. Patil, M. Beck, K. Pöppel, J. Brandstetter, G. Klambauer, R. Pascanu, S. Hochreiter: *A large recurrent action model: xLSTM enables fast inference for robotics tasks*, 2024. URL https://arxiv.org/abs/2410.22391.

xLSTM Paper

M. Beck, K. Pöppel, M. Spanring, A. Auer, O. Prudnikova, M. Kopp, G. Klambauer, J. Brandstetter, S. Hochreiter: »xLSTM: Extended long short-term memory«. In: *Proceedings of the Conference on Neural Information Processing Systems (NeurIPS)*, 2024. URL https://arxiv.org/abs/2405.04517.

Zu Abstraktion von LSTM

Y. Lakretz, G. Kruszewski, T. Desbordes, D. Hupkes, S. Dehaene, and M. Baroni: »The emergence of number and syntax units in LSTM language models«.

In: J. Burstein, C. Doran, and T. Solorio (eds.), *Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 11–20. Association for Computational Linguistics, 2019. doi: 10.18653/v1/N19-1002.

Zu LSTM für Flutvorhersagen und Google-App, Nature-Publikation

G. Nearing, D. Cohen, V. Dube, M. Gauch, O. Gilon, S. Harrigan, A. Hassidim, D. Klotz, F. Kratzert, A. Metzger, S. Nevo, F. Pappenberger, C. Prudhomme, G. Shalev, S. Shenzis, T. Y. Tekalign, D. Weitzner, and Y. M. B. Kosko: *Global prediction of extreme floods in ungauged watersheds.* Nature, 627:559–563, 2024. doi: 10.1038/s41586-024-07145-1.

Zu LSTM in der Hydrologie und Flutvorhersagen

F. Kratzert, D. Klotz, C. Brenner, K. Schulz, and M. Herrnegger: *Rainfall-runoff modelling using long short-term memory (LSTM) networks.* Hydrology and Earth System Sciences, 22(11):6005–6022, 2018. F. Kratzert, D. Klotz, G. Shalev, G. Klambauer, S. Hochreiter, and G. Nearing: *Benchmarking a catchment-aware long short-term memory network (LSTM) for large-scale hydrological modeling.* ArXiv, 1907.08456, 2019.

Zu Matrix Memory of xLSTM

T. Kohonen: *Correlation matrix memories.* IEEE Transactions on Computers, C-21(4), 1972. doi: 10.1109/TC.1972.5008975.

Zu Matrix Memory von xLSTM

K. Nakano: *Associatron – a model of associative memory.* IEEE Transactions on Systems, Man, and Cybernetics, SMC-2(3):380–388, 1972. doi: 10.1109/TSMC.1972.4309133.

Zu modernen Hopfield-Netzwerken

D. Krotov and J. J. Hopfield: »Dense associative memory for pattern recognition«. In: D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, pp. 1172–1180. Curran Associates, Inc., 2016.

Zusammenhang Transformer und LSTM

A. Katharopoulos: A. Vyas, N. Pappas, and F. Fleuret. Transformers are RNNs: »Fast autoregressive transformers with linear attention«. In: E. H. Daumé III and A. Singh (eds.), *International Conference on Machine*

Learning (ICML), volume 119 of Proceedings of Machine Learning Research, pp. 5156–5165. PMLR, 2020.

Zusammenhang Transformer zu linearen Ansätzen / Schmidhuber

I. Schlag, K. Irie, and J. Schmidhuber: »Linear transformers are secretly fast weight programmers«. In: M. Meila and T. Zhang (eds.), *Proceedings of the 38th International Conference on Machine Learning (ICML)*, volume 139 of Proceedings of Machine Learning Research, pp. 9355–9366. PMLR, 2021.

Zusammenhang von Transformern zu modernen Hopfield-Netzwerken

H. Ramsauer, B. Schäfl, J. Lehner, P. Seidl, M. Widrich, L. Gruber, M. Holzleitner, M. Pavlovic, G. K. Sandve, V. Greiff, D. Kreil, M. Kopp, G. Klambauer, J. Brandstetter, and S. Hochreiter: »Hopfield networks is all you need«. In: *International Conference on Learning Representations (ICLR)*. OpenReview, 2021.

Über den Autor



Sepp Hochreiter, geboren 1967 in Mühldorf am Inn, ist einer der führenden Wissenschaftler und Pioniere im Bereich der künstlichen Intelligenz. Er entwickelte die LSTM-Technologie (Long Short-Term Memory), die maßgeblich zur Entwicklung von künstlicher Intelligenz beigetragen hat und zum Beispiel in Spracherkennungs- und Textverarbeitungstools wie Siri, Alexa oder Google Translate verwendet wird.

Seit 2006 ist er Universitätsprofessor an der Johannes Kepler-Universität in Linz, wo er das Institute for Machine Learning und das Labor für Artificial Intelligence leitet. 2023 hat er die KI-Firma NXAI gegründet, bei der er als Chief Scientist tätig ist.

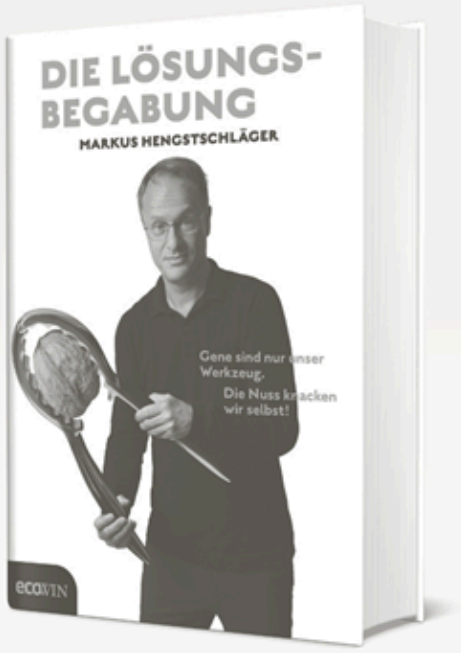
GIBT ES LEBEN ODER GAR EINE ZWEITE ERDE IN UNSERER GALAXIE?



Die Frage nach der Existenz außerirdischen Lebens beschäftigt die Menschheit seit langem. Lisa Kaltenegger, die international gefragte Astrophysikerin und Astronomin, begibt sich in diesem Buch auf Spurensuche nach Super-Erden, Felsplaneten und Habitablen Zonen. Sie präsentiert Ergebnisse aktueller Forschung auf unterhaltsame Weise und erklärt, unter welchen Umständen außerirdisches Leben möglich ist.

LISA KALTENEGGER
SIND WIR ALLEIN IM UNIVERSUM?
208 Seiten · 14,5 × 21,0 cm
Hardcover
978-3-7110-0350-8 · € 24,00

DIE LÖSUNGSBEGABUNG DES MENSCHEN ALS ANTWORT AUF HERAUSFORDERUNGEN



In einer zunehmend komplexen Zeit des Wandels, der digitalen Revolution und der politischen sowie klimatischen Krisen ist die Lösungsbegabung des Menschen mehr denn je gefragt. Der vielfach ausgezeichnete Genetiker Markus Hengstschläger plädiert in diesem Buch dafür, die Innovationskraft des Menschen zu steigern. Die gleichzeitige Verfolgung gerichteter sicherer und ungerichteter flexibler Strategien für mehr und für weniger vorhersehbare Zukünfte schafft den notwendigen Mut aus Sicherheit, um die alten Wege zu verlassen. Angewendet auf Bildung, Talentmanagement, Forschung, Politik und unsere Arbeitswelt können wir dadurch die Geschichte der Gegenwart neu schreiben.

MARKUS HENGSTSCHLÄGER
DIE LÖSUNGSBEGABUNG
256 Seiten · 14,5 × 21,0 cm
Hardcover mit Schutzumschlag
978-3-7110-0279-2 · € 24,00



Kreisky, Israel und die Juden

Aschheim, Daniel
9783711053800
208

Titel jetzt kaufen und lesen (Werbung)

Eine neue Perspektive auf das Wirken des langjährigen österreichischen Bundeskanzlers Bruno Kreisky

Bruno Kreisky wird in Österreich über alle Parteigrenzen hinweg bewundert, ja verehrt. Ganz anders in Israel: Hier gilt er als umstritten, wenn nicht gar als Feindbild. Gerade wegen seiner jüdischen Herkunft wurde ihm seine oft scharfe Kritik an der israelischen Politik und seine frühe Unterstützung der Palästinensischen Befreiungsorganisation PLO übel genommen.

Der israelische Diplomat Dr. Daniel Aschheim und Autor zeichnet in diesem Buch ein facettenreiches Bild von dem Staatsmann, der die Zweistaatenlösung in die Friedensbemühungen um den Nahostkonflikt einbrachte.

- Biografie: Kreiskys Persönlichkeit im Spiegel von Geschichte und Gegenwart des Judentums in Österreich

- Interviews, Archivmaterial und Einblicke in die israelische Diplomatie der 1970er Jahre zeichnen ein differenziertes Bild des Kultkanzlers

- Friedenspolitik für Israel und Palästina: Was bedeutete Kreiskys jüdische Identität für sein politisches Handeln?

- Kompliziertes Beziehungsgeflecht: Kreisky, das Judentum und der Staat Israel

- Politisches Sachbuch über die Ära Kreisky in Österreichs Zweiter Republik

Ein Stück Zeitgeschichte: Die österreichische Politik und der Jahrhundert-Kanzler

Bruno Kreisky zählt zweifellos zu den berühmtesten Persönlichkeiten Österreichs: In seiner 13-jährigen Amtszeit als Bundeskanzler prägte er das politische Nachkriegs-Österreich wie kein anderer. Daniel Aschheim zeichnet in seinem Buch ein differenziertes Bild des Ausnahme-Politikers und stellt uns einen Mann voller Visionen, aber auch voller Widersprüche vor. Ungeachtet der eigenen Haltung zu Person und Wirken Kreiskys ist eines klar: Dieser Mann lässt niemanden kalt.

Begleiten Sie den Autor auf eine Reise durch die politische Zeitgeschichte Österreichs in der Nachkriegszeit!

Titel jetzt kaufen und lesen (Werbung)



Mit Zuversicht

Schüssel, Wolfgang
9783711053671
312

Titel jetzt kaufen und lesen (Werbung)

Inspirierende Geschichte(n)

Unsere Gegenwart ist geprägt von ständiger Veränderung, oft begleitet von sozialer Ungerechtigkeit und politischer Instabilität. Viele Menschen fühlen sich unsicher, sind unzufrieden und suchen

nach Hoffnung und Zuversicht. Wolfgang Schüssel wirft einen Blick zurück auf bewegende Momente in der Geschichte sowie auf einschneidende Augenblicke seiner politischen Karriere – und zeigt, dass wir auch in schwierigen Zeiten zuversichtlich sein dürfen.

Der ehemalige österreichische Bundeskanzler Wolfgang Schüssel feiert seinen 80. Geburtstag mit einem besonderen Werk: Er schreibt Geschichten über Zuversicht und Mut. Er versammelt Erzählungen über wegweisende Persönlichkeiten der Vergangenheit aus Politik und Kultur wie Konrad Adenauer, Viktor Frankl, Wolfgang Schäuble oder Manès Sperber, Geschichten aus der Geschichte und Anekdoten aus seinem privaten Leben. Es sind Erzählungen, die uns zeigen, wie wir aus den Lehren der Vergangenheit Kraft schöpfen können und was wir von gestern für heute lernen können. Es sind Geschichten, die ermutigen und den Glauben an eine bessere Zukunft stärken.

Titel jetzt kaufen und lesen (Werbung)



Hausverstand

Doskozil, Hans Peter
9783711053381
192

Titel jetzt kaufen und lesen (Werbung)

Stationen der Lebensgeschichte eines Sozialdemokraten, der polarisiert

Hans Peter Doskozil ist ein Politiker aus Überzeugung. Der burgenländische Landeshauptmann hat sich ganz der Verbesserung

der Lebensumstände in "seinem" Bundesland verschrieben. Dass er dabei auch aneckt und selbst in der eigenen Partei nicht nur Unterstützer hat, gehört für ihn zum Geschäft.

In diesem Buch zeichnet er seine Lebensgeschichte, verwoben mit seinem beruflichen Werdegang, auf höchst spannende Weise nach. Eine Zeitreise von der Kindheit in Kroisegg bis ins Jahr 2024.

- Politische Werte im Sinne der Sozialdemokratie und was der Hausverstand damit zu tun hat

- Biografie eines Politikers: Wie der streitbare Burgenländer mit Schicksalsschlägen und Niederlagen umgeht

- Hans Peter Doskozils politische Heimat: Verbundenheit und Konflikte mit der SPÖ

- Politik in Österreich: Was der Landeshauptmann in den nächsten Jahren für sein Burgenland plant

- Eurofighter, Mindestlohn, Migration und Integration: Stationen einer politischen Karriere

Mit Sach- und Hausverstand gegen Machtmissbrauch und Ungerechtigkeit

Hans Peter Doskozil hat mit dem Kapitel Bundespolitik abgeschlossen und sich ganz dem Burgenland verschrieben. Was er hier bereits erreicht hat und was in den nächsten Jahren ansteht, verrät er in diesem biografischen Buch auf sehr persönliche Weise. Seine Politik, seine Überzeugungen und wie er gelernt hat, mit seiner Kehlkopferkrankung zu leben, erzählt er in dieser

Lebensgeschichte, die vor allem eines zeigt: Dass der streitbare SPÖ-Politiker sich ganz in den Dienst seiner Heimat stellt – mit Hausverstand.

Titel jetzt kaufen und lesen (Werbung)



Anpacken und anpassen

Zehetner, Elisabeth

9783711053831

176

Titel jetzt kaufen und lesen (Werbung)

Klimafit in die Zukunft

Begrünte Fassaden, Wasserspeicher, innovative Landwirtschaft – all das kann uns helfen, den Klimawandel aufzuhalten und seine Auswirkungen abzumildern. Elisabeth Zehetner zeigt, wie auch die Wirtschaft durch Klimainnovationen profitieren kann.

Unsere Fähigkeit zur Anpassung und die Entwicklung innovativer Lösungen sind entscheidend, um den Herausforderungen des Klimawandels erfolgreich zu begegnen. Durch kluge Stadtplanung, fortschrittliche landwirtschaftliche Methoden, eine effiziente Nutzung von Wasserressourcen, präventive Maßnahmen gegen Naturkatastrophen und wirtschaftliche Anpassungsstrategien können wir den Auswirkungen des Klimawandels gezielt entgegenwirken und eine positive Zukunft für kommende Generationen gestalten. Aus diesem Grund richtet dieses Buch den Fokus auf die Anpassung an das bereits veränderte Klima, die genauso bedeutend ist wie ein verantwortungsvoller Klimaschutz.

Titel jetzt kaufen und lesen (Werbung)



Die Macht der Kränkung

Haller, Reinhard
9783711051394
248

Titel jetzt kaufen und lesen (Werbung)

An seelischen Verletzungen wachsen: So können Sie Kränkungen überwinden!

Kränkungen machen krank. Verletzte Gefühle und emotionale Übergriffe können unsere Selbstachtung nachhaltig erschüttern. Sie

zerstören uns selbst und unsere Beziehungen zu anderen. Häufig sind sie der Hauptauslöser für psychische Probleme wie Sucht, Depressionen oder Angststörungen. Doch niemand bleibt im Leben von Demütigungen und Beleidigungen verschont. Der renommierte Psychologe Dr. Reinhard Haller klärt über das Wesen der Kränkung und ihre Wirkungsmechanismen auf. Er gibt konkrete Tipps, wie wir eigene seelische Verletzungen verarbeiten können, und zeigt, wie wir zu einem achtsamen und empathischen Umgang mit unseren Mitmenschen gelangen.

- Der wunde Punkt: Was trifft uns und wie können wir unsere Persönlichkeit stärken?

- Enttäuschung, Ablehnung, Mobbing: wie seelische Verletzungen zu psychischen Erkrankungen führen

- Der richtige Umgang mit negativen Gefühlen: so können wir Kränkungen überwinden

- Vom gefragten Kriminalpsychiater und Gerichtsgutachter Prof. Dr. Reinhard Haller

- Mit bewegenden Fallbeispielen aus der (Kriminal-)Geschichte

Am Anfang stand die Kränkung… Das Zerstörungspotenzial psychischer Gewalt

Die spektakulärsten Auswirkungen von Kränkungen zeigen sich in der Kriminalgeschichte: Kaum ein Mörder, Terrorist oder Attentäter, dessen Motive nicht in vorangegangener Demütigung, beispielsweise als Mobbing-Opfer, zu finden sind. Wie seelische

Verletzungen und ein daraus resultierendes geringes Selbstwertgefühl zu einem Amoklauf führen können, zeigt die Fernsehserie, die auf Basis dieses Sachbuchs entstand: "Am Anschlag. Die Macht der Kränkung" (Ö/D 2021). Der Arzt und Psychotherapeut Reinhard Haller erklärt anschaulich, wie sehr Kränkungen unsere Selbstachtung angreifen können. Doch mit seinen erprobten Tipps gelingt es, an seelischen Verletzungen nicht nur zu wachsen, sondern auch die eigene Persönlichkeit zu stärken!

Titel jetzt kaufen und lesen (Werbung)