„KI hat eigentlich keinen blassen Schimmer, was sie tut“

KI verändert die Wissenschaft – auch die Neurowissenschaft. Wir haben mit Gabriele Lohmann über Forschung an der Schnittstelle von Gehirn und Algorithmus gesprochen: über künstliche Intelligenz als Forschungswerkzeug, das Decodieren von Hirnaktivität, und warum große Sprachmodelle im Grunde nicht verstehen, was sie tun.

Zum einen ist die Idee hinter den aktuellen KI-Modellen – künstliche neuronale Netzwerke – ja dem Aufbau des menschlichen Gehirns nachempfunden. Neurale Netze gibt es schon seit den 1950er Jahren; damals war man inspiriert von Neuronen und ihren Verbindungen. Natürlich gibt es trotzdem große Unterschiede zwischen Gehirnen und künstlichen neuronalen Netzen. Aber KI und Neurowissenschaft befruchten sich gegenseitig: Die KI lernt vom Gehirn, und wir Neurowissenschaftler profitieren umgekehrt von KI-Systemen, die uns bei der Arbeit helfen.

Wie genau hilft KI aktuell in der neurowissenschaftlichen Forschung?

Zum Beispiel bei der Bildverarbeitung, einem Standardproblem in der Magnetresonanztomographie (MRT). Wir erzeugen Bilder vom Gehirn, die ausgewertet werden müssen. Automatische Bildverarbeitungsprogramme dafür gibt es schon seit etwa 40 Jahren, aber sie haben nie wirklich stabil funktioniert. Erst im letzten Jahrzehnt hat die KI-Forschung hier einen gigantischen Sprung gemacht. Bemerkbar macht sich das zum Beispiel in der Alzheimer-Diagnostik: Man bestimmt dazu die Größe des Hippocampus auf einem MRT-Bild. Das ist eine der etabliertesten Marker der Erkrankung, denn der Hippocampus gehört zu den ersten Hirnregionen, die bei Alzheimer schrumpfen. Die manuelle Segmentierung – also das Einzeichnen des Hippocampus von Hand – ist extrem zeitaufwendig. KI-Tools schaffen das inzwischen zuverlässig und in kürzester Zeit.

Aber wie können wir überhaupt sicher sein, dass das, was die KI erkennt, der Realität entspricht?

Bei MRT-Aufnahmen ist das gut möglich, da man das Bild ja direkt vor sich hat; der Radiologe kann das Ergebnis schnell überprüfen. Die KI dient hier im Grunde als Hilfsmittel zur Beschleunigung.

Wirft der KI-Einsatz nicht grundsätzlich das Problem auf, dass man fürs Training sehr große Datenmengen braucht – womöglich mehr Daten, als zur Verfügung stehen?

Das kommt in der Tat vor. Ein aktuelles Projekt aus unserer Forschungsgruppe: Einer meiner Doktoranden beschäftigt sich mit der erwähnten Alzheimer-Früherkennung. Viele ältere Menschen zeigen leichte kognitive Einschränkungen, die noch keine Alzheimer-Diagnose rechtfertigen, sich aber zu einer solchen weiterentwickeln könnten. Um das vorherzusagen, braucht man hochgenaue Segmentierungen von MRT-Bildern. Für 3-Tesla-MRT-Geräte, wie sie in vielen Kliniken stehen, gibt es große Datenbanken mit tausenden von Bildern aus aller Welt; das reicht gut für das Training einer KI. Wir wollten das auf Daten von unserem MRT-Gerät mit 9,4-Tesla Feldstärke anwenden, aber solche Hochfeld-Magnetresonanztomographen finden sich sehr selten; wir hatten daher nur Aufnahmen von etwa 100 Versuchspersonen aus unserem Institut zur Verfügung. Die Lösung meines Doktoranden war, das auf 3-Tesla-Daten trainierte Modell als Ausgangspunkt zu nehmen und es für 9,4 Tesla anzupassen. Und es funktioniert.

Gibt es bereits Ansätze, bei denen KI in der Forschung selbstständig Muster aus großen Datenmengen extrahiert und daraus eigene Hypothesen generiert?

Ja, in gewisser Weise schon, beim sogenannten Decoding: Man misst die Hirnaktivität von Probanden im Scanner und versucht zu rekonstruieren, woran sie gerade denken. Das klappt aktuell so leidlich – mithilfe von sogenannten Diffusionsmodellen, der Technologie hinter Deepfakes. Man legt einen Probanden in den Scanner, zeigt ihm Bilder, und die KI versucht zu erraten, was er sieht. Auf Grundlage eines Basistrainings mit großen Datenmengen erkennt das Modell Muster; den Rest „lügt" es mithilfe

von Deepfake-Technologie dazu.

Bleibt das reine Grundlagenforschung, oder sind auch Anwendungen denkbar – etwa für Menschen mit Locked-in-Syndrom?

Langfristig könnte man sich genau das vorstellen. So weit sind wir aber noch nicht.

Wohin wird sich die Entwicklung in den nächsten fünf bis zehn Jahren bewegen?

Aktuell wird KI unglaublich viel als Werkzeug genutzt, zum Beispiel beim Schreiben von Artikeln, oder auch, um sich Vorschläge machen zu lassen: Man gibt der einem großen Sprachmodell wie ChatGPT ein Manuskript und fragt: Wie wichtig ist das? Was würdest du als nächstes tun? Die Vorschläge sind erstaunlich gut; wirklich verblüffend, wie klug das wirkt. Aber schaut man genauer hin, merkt man: Diese Modelle haben eigentlich keinen blassen Schimmer, was sie da tun. Im Grunde betreiben sie Mustererkennung: Sie verarbeiten riesige Datenmengen und berechnen statistisch, welches Wort wahrscheinlich als nächstes kommt. Echtes Verständnis steckt nicht dahinter. Mich interessiert deshalb gerade besonders, wie man testen kann, was große Sprachmodelle wirklich verstehen. Bestehende Tests werden regelmäßig nach etwa zwei Jahren geknackt, weil die Entwicklung so rasant voranschreitet. Ich versuche daher, einen neuen zu entwickeln, der auf logisches Denken abzielt. Bis eine KI das beherrscht, dürfte es noch eine Weile dauern.

Pressemitteilung

13.05.2026
Quelle: Max-Planck-Institut für Biologische Kybernetik

Weitere Informationen

Kontakt
PD Dr. Gabriele Lohmann
Gruppenleiterin
Tel.: +49 (0)7071 601 931
E-Mail: gabriele.lohmann(at)tuebingen.mpg.de

- Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.
- Max-Planck-Institut für biologische Kybernetik