

Künstliche Intelligenz kann psychiatrische Diagnosestellung verbessern

Ein Kölner Forschungsteam testete, wie drei große Sprachmodelle Überschneidungen und Redundanzen in klinischen Fragebögen zu psychischen Erkrankungen feststellen können / Veröffentlichung in „Nature Mental Health“

Große Sprachmodelle können dazu beitragen, die Fragebögen zur Diagnosestellung psychischer Erkrankungen zu verbessern, indem sie die Verallgemeinerbarkeit von Symptomen optimieren und Redundanzen reduzieren. Sie können sogar zu neuartigen Konzeptualisierungen psychischer Störungen beitragen. Das ist das Ergebnis einer internationalen Studie unter Leitung von Professor Dr. Joseph Kambeitz und Professor Dr. Kai Vogeley von der Medizinischen Fakultät der Universität zu Köln und der Uniklinik Köln. Der Artikel „The empirical structure of psychopathology is represented in large language models“ ist in der Fachzeitschrift Nature Mental Health erschienen.

Um eine psychische Erkrankung zu diagnostizieren, verlassen sich Mediziner*innen unter anderem auf die Symptome, die Patient*innen selbst berichten und die in klinischen Fragebögen festgehalten werden. In solchen Fragebögen ist oft die genaue sprachliche Formulierung einzelner Fragen entscheidend, um die richtige Diagnose zu stellen. Hierbei sind etablierte Fragebögen jedoch oft sehr variabel. Die Forschenden sehen Hinweise, dass es bei der Feststellung von Depression, bipolarer Störung und dem Risiko einer Psychose inhaltliche Überschneidungen in den Fragen gibt und auch Abweichungen, was eine genaue Diagnose erschwert.

Darüber hinaus verlassen sich Ärzt*innen auf ihre klinische Erfahrung. Das bedeutet, dass sie einzelne Symptome mit einer bestimmten Erkrankung assoziieren, die ihrer individuellen Erfahrung entspricht. Da verschiedene Krankheiten jedoch gleiche oder ähnliche Symptome hervorbringen können, kann dies ebenfalls das Risiko der Falschdiagnose erhöhen. „Wir wissen erstaunlich wenig darüber, ob und wie Formulierungen in klinischen Fragebögen bestimmte Assoziationen bei Ärztinnen und Ärzten auslösen“, sagt Professor Joseph Kambeitz. Inkonsistente Ergebnisse könnten zudem aus Unterschieden zwischen Patient*innen derselben Diagnosegruppe oder – alternativ – aus Unterschieden zwischen Fragebögen resultieren.

Ein Ansatz, die sprachvermittelten Krankheitsbeschreibungen zu untersuchen, bieten sogenannte große Sprachmodelle (Large Language Models, kurz LLMs). Das Team nutzte die LLMs GPT-3, Llama und BERT, um sowohl die Struktur als auch den Inhalt von vier klinischen Fragebögen zu untersuchen. Grundlage waren Daten aus über 50.000 Fragebögen zu Depression, Angst, Psychoserisiko und Autismus.

In der klinischen Praxis treten Symptome häufig parallel auf – wie etwa empirische Assoziationen von Antriebsmangel und Freudverlust. Die Analyse zeigte, dass die LLMs „wissen“, welche Symptome in der Praxis gleichzeitig auftreten. Ohne Zugang zu spezifischen empirischen Daten, zeigen sich die gleichen Symptomassoziationen in LLMs rein auf Basis der Fragebogenformulierungen.

Damit eröffnen sich neue Perspektiven, wie Künstliche Intelligenz künftig psychologische Fragebögen verbessern, redundante Items vermeiden und die Diagnostik sowie das Verständnis

psychischer Erkrankungen effizienter gestalten kann. LLMs können dazu dienen, Fragebögen zu entwickeln, die präzise sind, also psychologische Symptome sicher erkennen, und gleichzeitig effizient – sie stellen nur so viele Fragen wie nötig und entlasten somit Patient*innen und Behandelnde.

„KI kann nicht nur medizinisches Wissen abbilden, sondern auch die Strukturen psychischer Erkrankungen. Das ist ein wichtiger Schritt, um digitale Methoden und Neurowissenschaften enger zusammenzuführen und Diagnostik sowie Forschung in der Psychiatrie weiterzuentwickeln“, sagt Professor Kai Vogeley.

Professor Joseph Kambeitz resümiert: „In der Psychiatrie nimmt das ‚gesprochene Wort‘ in Diagnose und Therapie einen großen Stellenwert ein. Aktuell untersuchen zahlreiche vielversprechende Projekte, wie wir LLMs in der Psychiatrie anwenden können – von der Diagnostik über das Verfassen und Ergänzen von Befunden bis hin zur Simulation von Therapiegesprächen. Auf diesem Gebiet sind noch viele spannende Forschungsergebnisse zu erwarten.“

Originalpublikation:

<https://www.nature.com/articles/s44220-025-00527-y>

DOI: 10.1038/s44220-025-00527-y