



Ressort: Mixed News

Chemie-KIs müssen nichts von Chemie verstehen

Bonn, 16.10.2025 [ENA]

Chemie-KIs müssen nichts von Chemie verstehen.

Studie der Universität Bonn belegt, dass Transformer-KI-Modelle ausschließlich statistische Korrelationen lernen.

Sprachmodelle kommen inzwischen auch in den Naturwissenschaften zum Einsatz. In der Chemie werden sie zum Beispiel genutzt, um potenzielle Wirkstoffe vorherzusagen. Dazu muss man sie zunächst aufwändig trainieren. Doch Wissen um biochemische Zusammenhänge

eignen sie sich dabei wohl nicht unbedingt an. Stattdessen ziehen sie ihre Schlüsse auf Basis von Ähnlichkeiten und statistischen Korrelationen, wie eine aktuelle Studie der Universität Bonn zeigt. Die Ergebnisse sind nun in der Fachzeitschrift „Patterns“ erschienen.

Große Sprachmodelle sind oft verblüffend gut in dem, was sie tun - ob sie nun mathematische Beweise führen, Musik komponieren oder einen Werbeslogan entwerfen. Doch wie kommen sie zu ihren Ergebnissen? Verstehen sie tatsächlich, was eine Symphonie ausmacht oder einen guten Witz? Es ist gar nicht so einfach, darauf eine Antwort zu geben. „Alle Sprachmodelle sind eine Black Box“, betont Prof. Dr. Jürgen Bajorath.

„Es ist schwer, ihnen - metaphorisch gesprochen - in den Kopf zu schauen.“

Dennoch hat der Chemieinformatiker am Lamarr-Institut für maschinelles Lernen und künstliche Intelligenz der Universität Bonn genau das versucht. Und zwar für eine besondere Form von KI-Algorithmen - die chemischen Transformer-Modelle. Sie funktionieren im Prinzip ganz ähnlich wie ChatGPT, Google Gemini oder Elon Musks „Grok“. Diese wurden mit riesigen Mengen von Texten trainiert und haben so gelernt, selbst Sätze zu formulieren. Chemische Sprachmodelle basieren dagegen in der Regel auf deutlich weniger Daten. Sie eignen sich ihr Wissen anhand molekularer Repräsentationen und Beziehungen an, zum Beispiel den sogenannten SMILES-Strings.

Das sind Zeichenfolgen, die Moleküle und deren Struktur als eine Sequenz von Buchstaben und Symbolen darstellen.

Redaktioneller Programmdienst: European News Agency

Annette-Kolb-Str. 16
D-85055 Ingolstadt
Telefon: +49 (0) 841-951. 99.660
Telefax: +49 (0) 841-951. 99.661
Email: contact@european-news-agency.com
Internet: european-news-agency.com

Haftungsausschluss:

Der Herausgeber übernimmt keine Haftung für die Richtigkeit oder Vollständigkeit der veröffentlichten Meldung, sondern stellt lediglich den Speicherplatz für die Bereitstellung und den Zugriff auf Inhalte Dritter zur Verfügung. Für den Inhalt der Meldung ist der allein jeweilige Autor verantwortlich.



International Press Service.....

- Trainingsdaten gezielt manipuliert -

In der Pharmazie sucht man oft nach Substanzen, die bestimmte Enzyme hemmen oder Rezeptoren blockieren. Mit chemischen Sprachmodellen kann man zum Beispiel versuchen, aktive Moleküle auf der Basis der Aminosäure-Sequenzen von Zielproteinen vorherzusagen. „Wir haben dieses sogenannte sequenzbasierte Moleküldesign als ein Testsystem benutzt, um besser zu verstehen, wie die Transformer zu ihren Vorhersagen kommen“, erklärt Bajoraths Doktorand Jannik Roth, der die Studie mit ihm durchgeführt hat.

„Wenn man ein solches Modell nach einer Anlernphase etwa mit einem neuen Enzym füttert, spuckt es mit etwas Glück eine Verbindung aus, mit der sich dieses Enzym hemmen lässt. Doch heißt das, dass die KI gelernt hat, nach welchen biochemischen Prinzipien so eine Hemmung erfolgt?“

In der Anlernphase werden chemische Sprachmodelle mit Paaren aus Aminosäure-Sequenzen und den dazugehörigen Wirkstoff-Molekülen trainiert. Um ihrer Frage nachzugehen, haben die Wissenschaftler diese Trainingsdaten gezielt manipuliert. „Wir haben das Modell beispielsweise zunächst nur mit einer bestimmten Gruppe von Enzymen und deren Hemmstoffen gefüttert“, erklärt Bajorath.

„Wenn wir nun für Testzwecke ein neues Enzym aus derselben Familie genutzt haben, schlug uns der Algorithmus tatsächlich einen plausiblen Hemmstoff vor.“ Anders sah es dagegen aus, wenn die Forscher im Test ein Enzym aus einer anderen Familie verwandten - also eines, das im Körper ganz andere Aufgaben übernimmt. Dann lieferte das chemische Sprachmodell nur unbrauchbare Ergebnisse.

- Statistische Daumenregel -

„Das spricht dagegen, dass das Modell allgemein gültige chemische Prinzipien erlernt hat - dass es also weiß, wie die Hemmung eines Enzyms chemisch in aller Regel abläuft“, sagt der Wissenschaftler. Stattdessen basieren die Vorschläge allein auf statistischen Zusammenhängen, also auf Mustern in den Daten.

Etwa: Wenn das neue Enzym einem der Enzyme ähnelt, die ich in der Trainingsphase gelernt habe, dann wird bei ihm wohl auch ein ähnlicher Hemmstoff funktionieren. „Eine solche Daumenregel muss nicht unbedingt schlecht sein“, betont Bajorath, der am Lamarr-Institut den Bereich KI in den Lebenswissenschaften und der Medizin verantwortet. „Schließlich kann sie beispielsweise dabei helfen, neue Einsatzgebiete für bekannte Wirkstoffe zu finden.“

Allerdings ließen die Modelle in der Studie auch bei der Abschätzung von Ähnlichkeiten jede biochemische

Redaktioneller Programmdienst: European News Agency

Annette-Kolb-Str. 16
D-85055 Ingolstadt
Telefon: +49 (0) 841-951. 99.660
Telefax: +49 (0) 841-951. 99.661
Email: contact@european-news-agency.com
Internet: european-news-agency.com

Haftungsausschluss:

Der Herausgeber übernimmt keine Haftung für die Richtigkeit oder Vollständigkeit der veröffentlichten Meldung, sondern stellt lediglich den Speicherplatz für die Bereitstellung und den Zugriff auf Inhalte Dritter zur Verfügung. Für den Inhalt der Meldung ist der allein jeweilige Autor verantwortlich.



..... International Press Service.....

Kenntnis vermissen. So hielten sie Enzyme (oder auch Rezeptoren und andere Proteine) stets dann für ähnlich (und schlugen dementsprechend auch ähnliche Hemmstoffe vor),

wenn diese in 50 bis 60 Prozent ihrer Aminosäure-Sequenz übereinstimmten. Den Rest der Sequenzen konnten die Forscher beliebig durcheinander würfeln. Allerdings sind oft nur ganz bestimmte Teile eines Enzyms nötig, damit es seine Aufgabe erledigen kann. Eine einzige geänderte Aminosäure an dieser Stelle kann es funktionsunfähig machen. Andere Bereiche sind dagegen für die Funktion weniger relevant. „Die Modelle lernten während ihres Trainings nicht, wichtige von unwichtigen Sequenzanteilen zu unterscheiden“, betont Bajorath.

- Modelle plappern zuvor Gehörtes nach -

Die Ergebnisse der Studie zeigen daher eindrucksvoll, dass den Modellen zumindest für dieses Testsystem jegliches tiefere chemische Verständnis abgeht. Bildlich ausgedrückt, plappern sie nur mit leichten Variationen nach, was sie irgendwann in einem ähnlichen Kontext aufgeschnappt haben. „Das bedeutet nicht, dass sie für die Wirkstoff-Forschung ungeeignet sind“, betont Bajorath, der auch Mitglied im Transdisziplinären Forschungsbereich (TRA) „Modelling“ der Universität Bonn ist. „Es ist gut möglich, dass sie Wirkstoffe vorschlagen, die tatsächlich bestimmte Rezeptoren blockieren oder Enzyme hemmen.

Das tun sie aber mit Sicherheit nicht, weil sie die Chemie so gut verstehen, sondern weil sie Ähnlichkeiten in textbasierten molekularen Repräsentationen und statistischen Korrelationen erkennen, die uns verborgen bleiben. Das diskreditiert ihre Ergebnisse nicht. Man darf sie aber auch nicht überinterpretieren.“

Beteiligte Institutionen und Förderung:

Die Arbeiten wurden durch die Studienstiftung des deutschen Volkes finanziell gefördert.

Publikation: Jannik P. Roth, Jürgen Bajorath: Unraveling learning characteristics of transformer models for molecular design, Patterns, <https://doi.org/10.1016/j.patter.2025.101392>, URL: [https://www.cell.com/patterns/fulltext/S2666-3899\(25\)00240-5](https://www.cell.com/patterns/fulltext/S2666-3899(25)00240-5)

Bericht online lesen:

https://www.european-news-agency.de/mixed_news/chemie_kis_muessen_nichts_von_chemie_verstehen-92368/

Redaktioneller Programmdienst: European News Agency

Annette-Kolb-Str. 16
D-85055 Ingolstadt
Telefon: +49 (0) 841-951. 99.660
Telefax: +49 (0) 841-951. 99.661
Email: contact@european-news-agency.com
Internet: european-news-agency.com

Haftungsausschluss:

Der Herausgeber übernimmt keine Haftung für die Richtigkeit oder Vollständigkeit der veröffentlichten Meldung, sondern stellt lediglich den Speicherplatz für die Bereitstellung und den Zugriff auf Inhalte Dritter zur Verfügung. Für den Inhalt der Meldung ist der allein jeweilige Autor verantwortlich.



..... International Press Service.....

Redaktion und Verantwortlichkeit:

V.i.S.d.P. und gem. § 6 MDStV: Wilhelm Fussel

**Redaktioneller Programmdienst:
European News Agency**

Annette-Kolb-Str. 16
D-85055 Ingolstadt
Telefon: +49 (0) 841-951. 99.660
Telefax: +49 (0) 841-951. 99.661
Email: contact@european-news-agency.com
Internet: european-news-agency.com

Haftungsausschluss:

Der Herausgeber übernimmt keine Haftung für die Richtigkeit oder Vollständigkeit der veröffentlichten Meldung, sondern stellt lediglich den Speicherplatz für die Bereitstellung und den Zugriff auf Inhalte Dritter zur Verfügung. Für den Inhalt der Meldung ist der allein jeweilige Autor verantwortlich.