

Künstliche Intelligenz lüftet Geheimnis von Millionen von Proteinen

Ein Forschungsteam der Universität Basel hat eine wahre Fundgrube von bislang nicht beschriebenen Proteinen entdeckt. Dank der jüngsten Entwicklungen im «Deep Learning» gelang es den Forschenden, Hunderte neue Proteinfamilien und sogar eine neue Art der Proteinfaltung aufzudecken. Die Studie ist jetzt in «Nature» veröffentlicht.

In den vergangenen Jahren hat «AlphaFold» die Proteinforschung revolutioniert. Die Künstliche Intelligenz (KI) wurde mit Daten von Proteinen gefüttert, die Wissenschaftlerinnen und Wissenschaftler in mehr als 50 Jahren zusammengetragen haben. Durch dieses Training ist «AlphaFold» nun in der Lage, die dreidimensionale Gestalt von Proteinen mit einer sehr hohen Genauigkeit vorherzusagen. Im letzten Jahr konnte man mithilfe der KI die Strukturen von etwa 215 Millionen Proteinen prognostizieren. Dies erlaubt Einblicke in die Struktur nahezu aller Proteine. Interessant ist dies vor allem für Proteine, die bisher noch nicht genauer experimentell untersucht wurden – ein aufwändiger und langwieriger Prozess.

«Es gibt inzwischen sehr viele Quellen, in denen man Informationen zu Proteinen finden kann. Sie liefern wertvolle Erkenntnisse darüber, wie sich Proteine entwickeln und wie sie funktionieren», sagt Joana Pereira, Leiterin der Studie. Dennoch stand die Forschung lange Zeit vor einem Daten-Dschungel. Dem Forschungsteam von Prof. Dr. Torsten Schwede am Biozentrum der Universität Basel und Forschungsleiter am Swiss Institute of Bioinformatics (SIB) ist es nun erstmals gelungen, einen Teil dieser verborgenen Informationen zu entschlüsseln.

Vogelperspektive offenbart neue Proteinfamilien und -faltungen

Die Forschenden haben ein interaktives Netzwerk bestehend aus 53 Millionen Proteinen, deren Struktur «AlphaFold» mit einer hohen Qualität modelliert hat, aufgebaut. «Dieses Netzwerk ist eine wertvolle Quelle, um im grossen Massstab unbekannte Proteinfamilien und deren Funktionen theoretisch vorherzusagen», unterstreicht Erstautorin Dr. Janani Durairaj. Auf diese Weise entdeckte das Team 290 neue Proteinfamilien und eine neue Art der Proteinfaltung. Sie kommt der Form einer Blume sehr nahe.

Auf Basis der Expertise der Forschungsgruppe von Torsten Schwede bei der Entwicklung und dem Unterhalt der Software SWISS-MODEL zur Proteinmodellierung, stellten sie das neu entwickelte Netzwerk als interaktive Plattform im Internet zur Verfügung. Diese wird als «Protein Universe Atlas» bezeichnet.

KI als wertvolles Werkzeug für die Forschung

Schliesslich setzten die Forschenden auf Deep Learning Tools, um in diesem Netzwerk nach Informationen zu suchen, die neue Wege für Innovationen in den Biowissenschaften und -technologie ebnen. «Um neue Wirkstoffe für Medikamente zu finden oder mit biotechnologischen Verfahren die Funktion von Proteinen zu verändern, müssen wir zuerst die Struktur von Proteinen kennen und verstehen, wie sie funktionieren», sagt Pereira.

Das Swiss Institute of Bioinformatics förderte das Forschungsprojekt mit einem «Kickstarter» Grant,

der ins Leben gerufen wurde, um den Einsatz von KI in den Biowissenschaften voranzubringen. Sie unterstreicht das transformative Potenzial von Deep Learning und intelligenten Algorithmen in der Forschung.

Mit dem «Protein Universe Atlas» können Wissenschaftlerinnen und Wissenschaftler weltweit nun mehr über die für ihre Forschung relevanten Proteine erfahren. «Wir hoffen, dass diese neue Plattform nicht nur Forschenden und Biokuratoren, sondern auch Studierenden und Lehrenden helfen wird, mehr über die Proteinviefalt zu lernen, von der Struktur über die Funktion bis hin zur Evolution», fügt Janani Durairaj hinzu.

Originalpublikation

Janani Durairaj, Andrew M. Waterhouse, Toomas Mets, Tetiana Brodiazhenko, Minhal Abdullah, Gabriel Studer, Gerardo Tauriello, Mehmet Akdel, Antonina Andreeva, Alex Bateman⁶, Tanel Tenson, Vasili Hauryliuk, Torsten Schwede, Joana Pereira

[Uncovering new families and folds in the natural protein universe.](#)

Nature (2023), doi: 10.1038/s41586-023-06622-3

Weiterführende Informationen

- [Forschungsgruppe Prof. Dr. Torsten Schwede](#)
- [Protein Universe Atlas](#)