

Proteinmodellierung neu gedacht: KI-System BioEmu könnte Weg für bessere Medikamente ebnen

Forschende der Freien Universität Berlin an zukunftsweisender Proteinmodellierung von Microsoft Research beteiligt

Einem Forschungsteam der Freien Universität Berlin und von Microsoft Research AI for Science ist ein bedeutender wissenschaftlicher Durchbruch in der Modellierung biologischer Proteine gelungen - die Studie dazu ist gerade (10. Juli) in der Fachzeitschrift Science erschienen. Die Studie „Scalable emulation of protein equilibrium ensembles with generative deep learning“ stellt BioEmu vor - ein generatives Deep-Learning-System, das Gleichgewichtsverhalten von Proteinen mit bisher unerreichter Präzision und Genauigkeit emuliert.

Da die Funktion vieler Proteine unmittelbar von ihrer dynamischen Struktur abhängt, eröffnet BioEmu neue Perspektiven für das Wirkstoffdesign und könnte künftig dazu beitragen, die Erfolgsquote von neuen Wirkstoffen in klinischen Studien zu erhöhen.

Das neue KI-System BioEmu ist in der Lage, tausende statistisch unabhängiger Proteinstrukturen pro Stunde auf einer einzigen Grafikkarte (GPU) zu generieren. „Damit werden die Kosten und den Zeitaufwand erheblich gesenkt, um funktionell relevante Strukturveränderungen zu untersuchen“, sagt Prof. Dr. Frank Noé, der eine Honorarprofessur an der Freien Universität Berlin innehat und Leiter der Studie ist. Durch die Integration von über 200 Millisekunden molekulardynamischer Simulationen mit experimentellen Daten kann BioEmu Struktur-Ensembles und thermodynamische Eigenschaften mit nahezu experimenteller Genauigkeit vorhersagen.

Besonders bemerkenswert ist die Fähigkeit des Systems, komplexe und biologisch relevante Strukturänderungen wie verborgene Bindungstaschen, Bewegungen ganzer Protein-Domänen oder lokale Entfaltungen zu erfassen. Selbst Stabilitätsveränderungen von Proteinen kann BioEmu mit einer Genauigkeit vorhersagen, die sich mit Laboranalysen messen lässt. „Damit bietet es einen skalierbaren Ansatz zur Modellierung von Proteinfunktionen im großen Stil - etwa auf genomischer Ebene“, betont zudem Prof. Dr. Cecilia Clementi, die eine Einstein-Professur an der Freien Universität Berlin innehat.

Die Technologie und der gesamte Quellcode von BioEmu sind frei verfügbar unter der MIT-Lizenz. Zusätzlich hat Microsoft Research den umfangreichen Datensatz veröffentlicht, der zur Entwicklung von BioEmu diente: Mit über 100 Millisekunden Simulationen auf tausende verschiedene Proteinsysteme verteilt, handelt es sich um die bislang größte öffentlich zugängliche Sammlung sequenzdiverser Proteinsimulationen.

Obwohl die Forschungsarbeiten vollständig bei Microsoft durchgeführt wurden, würdigt die Freie Universität Berlin die maßgeblichen Beiträge ihrer affilierten Forschenden. Die Leitung der Studie lag bei Prof. Dr. Frank Noé, Partner Research Manager bei Microsoft Research AI for Science in Berlin und Honorarprofessor an der FU Berlin. Wesentliche Beiträge lieferte auch Prof. Dr. Cecilia Clementi, Einstein-Professorin für Physik an der Freien Universität Berlin, die während der Projektlaufzeit als Gastwissenschaftlerin bei Microsoft tätig war.

Originalpublikation:

<https://www.science.org/doi/10.1126/science.adv9817>