

Intelligente Algorithmen für die Genomforschung

Plattform „Kipoi“ verbessert Zugang zu Machine-Learning-Modellen

Um herauszufinden, welche Gene für Krankheiten wie Krebs oder Diabetes verantwortlich sind, wird heute häufig auf Methoden des Maschinellen Lernens zurückgegriffen. Damit die klinische Forschung Zugang zu den aktuellsten Algorithmen erhält, hat Prof. Julien Gagneur von der Technischen Universität München (TUM) in Zusammenarbeit mit Wissenschaftlerinnen und Wissenschaftlern anderer Universitäten und Forschungsinstitute die Open-Access-Plattform „Kipoi“ aufgebaut.

Obwohl die Bedeutung von Methoden des Maschinellen Lernens in der Genomforschung in den letzten Jahren stetig gewachsen ist, mussten Forscherinnen und Forscher häufig auf veraltete Software zurückgreifen. Für die klinische Forschung waren die aktuellsten Modelle meist nicht zugänglich, da diese auf nichtöffentlichen Servern gespeichert wurden. Das soll sich mit einer neuen kostenlosen Open-Access-Plattform ändern: Kipoi ermöglicht den einfachen Austausch von Machine-Learning-Modellen im Bereich der Genomforschung. Geschaffen wurde die Plattform von Julien Gagneur, Assistenzprofessor für Computational Biology an der TUM in Zusammenarbeit mit Forscherinnen und Forschern der University of Cambridge, der Stanford University, des European Bioinformatics Institute (EMBL-EBI) und des European Molecular Biology Laboratory (EMBL).

„Kipoi bietet sehr spannende Möglichkeiten, einzelne Genome zu verstehen.“ – Julien Gagneur, Assistenzprofessor für Computational Biology

Trainierte Modelle offen verfügbar

„Was Kipoi besonders macht, ist dass man hier frei zugänglich bereits trainierte Machine-Learning-Modelle finden kann“, sagt Julien Gagneur. „Die Algorithmen, die wir hier bereitstellen, haben schon die wichtigsten Daten als Input bekommen und daran gelernt. Sie sind also sofort einsatzbereit, da der zeitraubende Prozess, sie auf Daten anzuwenden, bereits abgeschlossen ist,“ sagt Anshul Kundaje, Assistenzprofessor in Stanford. Mehr als 2.000 trainierte Modelle sind auf Kipoi momentan frei zugänglich. Wie die neue Plattform den Wissensaustausch in der Genomik-Community beschleunigen und damit die Genomforschung als Ganzes voranbringen wird, zeigen die Forscherinnen und Forscher in einer aktuell in Nature Biotechnology veröffentlichten Studie.

Schnelle Algorithmen und einfache Bedienung

Kipoi vereinfacht den Zugang zu bereits trainierten Modellen und ermöglicht es Forscherinnen und Forschern dadurch, das sogenannte Transferlernen durchzuführen: Ein Modell, das bereits mit einem bestimmten Datensatz trainiert wurde, kann eine ähnliche Aufgabe schneller lernen.

Zudem vereinfacht Kipoi den Prozess der Dateneingabe in die dort gespeicherten Modelle: Durch standardisierte Dateiformate und Software-Frameworks beschränkt sich die Installation und Ausführung eines Modells auf drei einfache Befehle. Die Plattform ist so auch für diejenigen leicht zu bedienen, die bisher keine Erfahrung mit Maschinellen Lernen hatten.

Zum Verständnis einzelner Genome

Da Kipoi speziell auf Modelle ausgerichtet ist, die Genotyp und Phänotyp miteinander verknüpfen, wird es mithilfe der neuen Plattform einfacher werden, genetische Ursachen für Krankheiten ausfindig zu machen: „Die Deep-Learning-Modelle, die wir mit Kipoi zur Verfügung stellen, basieren auf massiven genomischen Datensätzen,“ sagt Julien Gagneur. „Dies bietet sehr spannende Möglichkeiten, einzelne Genome zu verstehen. Zum Beispiel genetische Varianten, die Krankheiten verursachen, zu identifizieren oder Mutationen in Tumoren zu interpretieren.“

Wie groß der Beitrag der Plattform für die Genomforschung sein wird, hängt allerdings auch von der Genomik-Community ab. „Wir hoffen, dass in Zukunft mehr Forscherinnen und Forscher ihre Modelle in unsere Plattform einbringen werden,“ sagt Oliver Stegle, Teamleiter am EMBL. „Nur so können wir die Genomik-Analyse zugänglich gestalten und der Community eine breitere Palette an prädiktiven maschinellen Lernwerkzeugen zur Verfügung stellen.“

Avsec, Z.; Gagneur, J.; Kundaje, A.; Stegle, O. et al. (2019). The Kipoi repository accelerates community exchange and reuse of predictive models for genomics. Nature Biotechnology. Published online 28/05/19; DOI: [10.1038/s41587-019-0140-0](https://doi.org/10.1038/s41587-019-0140-0).

Mehr Informationen:

- [Zur Plattform Kipoi](#)
- [Zum Professorenprofil von Julien Gagneur](#)
- [Weitere aktuelle Meldung zur Verwendung künstlicher Intelligenz in der Proteomforschung](#)