

Intelligent algorithms for genome research

May 29 2019



Credit: CC0 Public Domain

Although the importance of machine learning methods in genome research has grown steadily in recent years, researchers have often had to resort to using obsolete software. Scientists in clinical research often did not have access to the most recent models. This will change with the new free open access repository: Kipoi enables an easy exchange of machine

learning models in the field of genome research. The repository was created by Julien Gagneur, Assistant Professor of Computational Biology at the TUM, in collaboration with researchers from the University of Cambridge, Stanford University, the European Bioinformatics Institute (EMBL-EBI) and the European Molecular Biology Laboratory (EMBL).

Trained models freely available

"What makes Kipoi special is that it provides [free access](#) to machine learning models that have already been trained," says Julien Gagneur.

"What we are doing with Kipoi is not just sharing data and software, but sharing models and algorithms that are already trained on the most [relevant data](#). These models are ready to use, because all the cumbersome work of applying them to data has already been done," says Anshul Kundaje, Assistant Professor at Stanford. More than 2,000 trained models are currently freely accessible on Kipoi. In a recent study published in *Nature Biotechnology*, the researchers show that the new repository will accelerate exchange in the genomics community and thereby advance [genome research](#).

Fast algorithms and easy operation

Because Kipoi simplifies access to already trained models, researchers can perform transfer learning. This means that a model that has already been trained with a particular dataset is capable of learning a similar task faster. Kipoi also simplifies the process of feeding data into the models stored there: Standardized file formats and software frameworks reduce the installation and execution of a [model](#) to three simple commands. Those who previously had no experience in machine learning can thus also easily use the repository.

Understanding individual genomes

As Kipoi is oriented towards models that link genotype and phenotype, the new platform will make it easier to identify genetic causes of disease: "Kipoi puts the latest deep learning models trained on massive genomics data at the fingertips of clinical researchers," says Julien Gagneur. "This provides very exciting opportunities to understand individual genomes, for instance to pinpoint genetic variants causing diseases or to interpret mutations occurring in tumors."

However, the extent of the platform's contribution to genomic research will also depend on the genomics community. "We hope that in the future more researchers will bring their models to our repository," says Oliver Stegle, team leader at the EMBL-EBI. "That is the only way we can make genomics analysis accessible and ultimately make a wider range of predictive machine learning tools available to the genomics community."

More information: Žiga Avsec et al, The Kipoi repository accelerates community exchange and reuse of predictive models for genomics, *Nature Biotechnology* (2019). [DOI: 10.1038/s41587-019-0140-0](https://doi.org/10.1038/s41587-019-0140-0)

Kipoi Repository: kipoi.org/

Provided by Technical University Munich

Citation: Intelligent algorithms for genome research (2019, May 29) retrieved 19 April 2024 from <https://phys.org/news/2019-05-intelligent-algorithms-genome.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is

provided for information purposes only.