Artificial intelligence boosts proteome research

May 29 2019

The authors Mathias Wilhelm, Tobias Schmidt and Siegfried Gessulat. Credit: Astrid Eckert / TUM

Using artificial intelligence, researchers at the Technical University of Munich (TUM) have succeeded in making the mass analysis of proteins

from any organism significantly faster than before and almost error-free. This new approach is set to spur a considerable change in the field of proteomics, as it can be applied in both basic and clinical research.

The genome of any organism contains the blueprints for thousands of proteins which control almost all the functions of life. Defective proteins lead to serious diseases, such as cancer, diabetes or dementia. Therefore, proteins are also the most important targets for drugs.

To better understand life processes and diseases and develop more appropriate therapies, it is necessary for as many proteins as possible to be analyzed simultaneously. At present, mass spectrometry is used in order to determine the type and quantity of proteins in a biological system. However, the current methods of data analysis continue to produce many mistakes.

A team at the Technical University of Munich led by bioinformatics scientist Mathias Wilhelm and biochemist Bernhard Küster, Professor of Proteomics and Bioanalytics at the Technical University of Munich, has now succeeded in using proteomic data to train a neural network in such a way that it is able to recognize proteins much more quickly and with almost no errors.

A solution to a serious problem

Mass spectrometers do not measure proteins directly. They analyze smaller parts consisting of amino acid sequences with up to 30 building blocks. The measured spectra of these chains are compared with databases in order to assign them to a specific protein. However, the evaluation software can only use part of the information that the spectra contain. Therefore, certain proteins are not recognized or are recognized incorrectly.

"This is a serious problem," explains Küster. The neural network developed by the TUM team uses all the information of the spectra for the process of identification. "We miss fewer proteins and make 100 times fewer mistakes," says Bernhard Küster.

Applicable to all organisms

"Prosit," as the researchers call the AI software, is "applicable to all organisms in the world, even if their proteomes have never been examined before," explains Mathias Wilhelm. "This enables research which was previously inconceivable."

With the help of 100 million mass spectra, the algorithm has been so extensively trained that it can be used for all common mass spectrometers without any additional training. "Our system is the global leader in this field," says Küster.

A market worth billions

Clinics, biotech companies, pharmaceutical companies and research institutes are using high-performance devices of this kind—the market is already worth billions. With Prosit, it will be possible to develop even more powerful instruments in the future. Researchers and physicians will also be better abled and faster at searching for biomarkers in patients' blood or urine, or in monitoring therapies for their effectiveness.

The researchers also have high hopes for fundamental research. "The method can be used to track down new regulatory mechanisms in cells," says Küster. "We hope to gain a considerable amount of knowledge here, which, in the medium and long term, will be reflected in the treatment of diseases suffered by humans, animals and plants."

Wilhelm also expects that "AI methods such as Prosit will soon change the field of proteomics , as they can be used in almost every area of protein research."

More information: Siegfried Gessulat et al, Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning, *Nature Methods* (2019). DOI: 10.1038/s41592-019-0426-7

Provided by Technical University Munich

Citation: Artificial intelligence boosts proteome research (2019, May 29) retrieved 26 April 2024 from https://phys.org/news/2019-05-artificial-intelligence-boosts-proteome.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.