Predicting the toxicity of chemicals with AI

July 18, 2024 | Ori Schipper
Topics: Pollutants

Researchers at Eawag and the Swiss Data Science Center have trained AI algorithms with a comprehensive ecotoxicological dataset. Now their machine learning models can predict how toxic chemicals are to fish.

Chemicals play an important role in our everyday lives, for example in the production of food, medicines and various everyday goods. Their impact on human health and the environment is closely monitored using various control mechanisms. For instance, the EU stipulates in the REACH regulation that fish toxicity tests must be carried out for all chemicals with a minimum annual production volume of 10 tonnes. These tests are expensive – and require an estimated 50,000 fish each year in Europe.

Scientists have been working for several decades on alternative methods that are cheaper and, above all, do not require the use of laboratory animals. Great hopes are pinned on computer-based methods that can predict the effects of chemicals on fish.

Promising predictive power of the models

The aquatic research institute Eawag and the Swiss Data Science Center (SDSC) have joined forces to cuarte a comprehensive ecotoxicological dataset, made available to the scientific community, to help develop and benchmark new AI algorithms in ecotoxicology. The dataset, called “ADORE”, consists of around 26,000 data points that describe the effects of almost 2,000 chemicals on 140 fish species. It includes as well a large set of characteristics of both chemicals and species.

As the researchers explain in their recently published scientific paper, the machine learning models are good at predicting the toxicity of chemicals. “The deviations observed are within the range of normal biological fluctuations,” say the two lead authors of the publication, Lilian Gasser, data scientist at the

SDSC, and Christoph Schür, postdoctoral researcher at Eawag. The researchers therefore consider the investigated methods to be “promising for the prediction of acute fish mortality”. And those methods could be used for other species groups, provided similar available data.

“However, there are still limitations that need to be taken into account,” the researchers state self-critically. Although the algorithms provide useful predictions on average, they are still substantially off in some cases for individual fish species. For example, they overestimate the toxicity of a chemical for certain fish species and underestimate it for other species. “Evidently, the models are mainly influenced by a few chemical properties and do not yet adequately capture species-specific sensitivities,” says Gasser.

A proper testing procedure leads to meaningful results

In their work, Gasser and Schür took into account the fact that the way in which the data is divided into a training dataset and a test dataset has a decisive influence for proper evaluation of the machine learning models. "It is essential that the algorithm is tested only on chemicals that are not present in the training set in order to show that it is able to identify chemical characteristics that are truly predictive of toxicity," both Gasser and Schür comment.

The future of chemical safety

According to Gasser and Schür and their co-authors, it is unlikely that machine learning models and artificial intelligence will soon make fish toxicity tests obsolete, but they are likely to help reduce them in the long term. The researchers believe these models will provide a more targeted assessment of chemical safety, which in future will include other biological factors in addition to physicochemical properties of the chemicals and mortality data.

For example, the model predictions could be combined with the evaluations of a series of other – animal-free – tests, which are currently being developed and validated at Eawag using different fish cell lines. For the development of such a highly informative chemical safety system, the researchers are encouraging close cooperation with the regulatory authorities so that the translation of research into practice can be jointly advanced.

Cover picture: Fish are often used in experiments. Machine learning could be an alternative to fish testing. (Photo: AdobeStock)

Original publication

Gasser, L.; Schür, C.; Perez-Cruz, F.; Schirmer, K.; Baity-Jesi, M. (2024) Machine learning-based prediction of fish acute mortality: implementation, interpretation, and regulatory relevance, *Environmental Science: Advances*, 3(8), 1124-1138, doi:10.1039/d4va00072b, Institutional Repository
Schür, C.; Gasser, L.; Perez-Cruz, F.; Schirmer, K.; Baity-Jesi, M. (2023) A benchmark dataset for machine learning in ecotoxicology, *Scientific Data*, 10(1), 718 (20 pp.), doi:10.1038/s41597-023-02612-2, Institutional Repository

Financing / Cooperations

Eawag Swiss Data Science Center (SDSC) ETH Zürich EPFL European Partnership for the Assessment of Risks from Chemicals (PARC) Horizon Europe

Related Links

Projekt «ML Tox»

Contact

Christoph Schuer
Postdoctoral Scientist
Tel. +41 58 765 5684
christoph.schuer@eawag.ch

Annette Ryser
Science editor
Tel. +41 58 765 6711
annette.ryser@eawag.ch

https://www.eawag.ch/en/info/portal/news/news-archive/archive-detail/predicting-the-toxicity-of-chemicals-with-ai