Presse-Information

Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

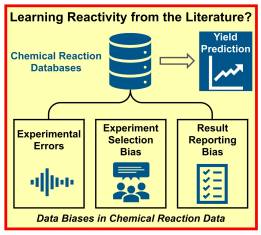
Boschstraße 12, 69469 Weinheim

Telefax: +49 (0) 6201/606-331
Telefon: +49 (0) 6201/606-315
E-Mail: angewandte@wiley-vch.de
www.angewandte.de
www.angewandte.org

Nr. 12/2022

More Data in Chemistry

Clearer reporting of negative experimental results would improve reaction planning in chemistry



Databases containing huge amounts of experimental data are available to researchers across a wide variety of chemical disciplines. However, a team of researchers have discovered that the available data is unsuccessful in predicting the yields of new syntheses using artificial intelligence (AI) and machine learning. Their study published in the journal *Angewandte Chemie* suggests that this is in large part down to the tendency of scientists not to report failed experiments.

Although AI-based models have been particularly successful in predicting molecular structures and material properties, they return rather inaccurate predictions for information relating to product yields in synthesis, as Frank Glorius and his team of researchers at Westfälische Wilhelms-Universität Münster, Germany, have discovered.

The researchers attribute this failure to the data used to train AI systems. “Interestingly, the prediction of reaction yields (reactivity) is much more challenging than the prediction of molecular properties. Reactants, reagents, quantities, conditions, the experimental execution—all determine the yield, and thus, the problem of yield prediction becomes very data-intensive,” explains Glorius. So, despite the huge amounts of available literature and results, the researchers came to realize that the data is not fit for accurate predictions of the expected yield.

Angewandte Chemie International Edition ist eine Wiley-VCH-Publikation / is a publication of Wiley-VCH.
Wiley-VCH GmbH - A company of John Wiley & Sons, Inc.
Sitz der Gesellschaft / Location of company: Weinheim.
Amtsgericht Mannheim, HRB 736569, Geschäftsführer: Sabine Steinbach, Dr. Guido F. Herrmann

The problem is not only down to a lack of experiments. In contrast, the team identified three possible causes for biased data. Firstly, the results of chemical syntheses may be flawed due to experimental error. Secondly, when chemists are planning their experiments, they may, either consciously or unconsciously, introduce bias based on personal experience and reliance on well-established methods. Finally, since only reactions with a positive outcome are believed to contribute to progress, failed reactions are reported less frequently.

To find out which of these three factors had the greatest influence, Glorius and the team purposely altered the datasets for four different, commonly used (and therefore data-rich) organic reactions. They artificially increased experimental error, reduced the size of the data sampling sets, or removed negative results from the data. Their investigations showed that the experimental error had the smallest influence on the model, while the contribution made by the lack of negative results was fundamental.

The group hopes that these findings will encourage scientists to always report failed experiments as well as their successes. This would improve data availability for training AI, ultimately helping to speed up planning and making experimentation more efficient. Glorius adds: “machine learning in (molecular) chemistry will increase efficiency dramatically and fewer reactions will have to be run to achieve a certain goal, for example, an optimization. This will empower chemists and will help them to make chemical processes—and the world—more sustainable.”

(3102 characters)

Frank Glorius, Westfälische Wilhelms-Universität Münster (Germany)
https://www.uni-muenster.de/Chemie.oc/glorius/glorius.html

Machine Learning for Chemical Reactivity: The Importance of Failed Experiments
Angewandte Chemie International Edition
doi: 10.1002/anie.202204647

based on can be found in our online pressroom at http://pressroom.angewandte.org.

Angewandte Chemie International Edition ist eine Wiley-VCH-Publikation / is a publication of Wiley-VCH.
Wiley-VCH GmbH - A company of John Wiley & Sons, Inc.
Sitz der Gesellschaft / Location of company: Weinheim.
Amtsgericht Mannheim, HRB 736569, Geschäftsführer: Sabine Steinbach, Dr. Guido F. Herrmann