JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

pubs.acs.org/jcim Editorial

# Editorial on Machine Learning

Cite This: J. Chem. Inf. Model. 2022, 62, 3941−3941 Read Online

ACCESS | Metrics & More | Article Recommendations

Downloaded via ZHEJIANG UNIV on October 7, 2022 at 07:03:14 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

Arguably, machine learning (ML), including deep learning (DL), is one of the most transformative technologies in human history. It is the main workhorse of data-driven discovery, the fourth paradigm of scientific discovery, after experiment, modeling, and simulation. ML has profoundly changed the landscape of scientific research in the past decade. The success of ML platforms, for example, AlphaFold2, has enabled data-driven breakthroughs in all areas of the chemical sciences, including new materials, biochemistry, and drug discovery.

The *Journal of Chemical Information and Modeling* (JCIM) welcomes manuscripts on a wide variety of ML-related topics, including the following:

- Introducing new ML methods to solve significant problems in the chemical sciences.
- Developing ML approaches that advance the chemical sciences.
- Developing new ML algorithms and software packages for the chemical sciences.
- Applications of existing techniques that shed light on the strengths and weaknesses of ML methods in the chemical sciences.
- Experimental and/or computational studies yielding new insight into the design of ML algorithms for the chemical sciences.
- Formalization of significant ML problems in the chemical sciences.
- Construction of new and significant ML datasets/databases in the chemical sciences.

The manuscript must clearly describe the status of the problem, its novelty, and its major and specific contributions. All methods and algorithms should be carefully validated and reproducible. If a manuscript introduces new terminology or a new ML technique that can be applied to chemical problems, it must justify why the current terminology or extant techniques are insufficient. Comparison with publicly available methods or models in the literature is required. Computer codes must be made available for independent validation and reproduction as required in our earlier editorial.

Data is a crucial component in ML approaches and applications. JCIM encourages experimental data-driven discovery in the chemical sciences. The use of any synthetic data or data generated from theoretical and/or mathematical models for ML is not encouraged in general, except for those that have been thoroughly validated. If a manuscript introduces new datasets, it must justify their significance within the chemical sciences. Datasets must be documented and made available for public use as required in our earlier editorial.

Manuscripts that apply existing methods to a single dataset and manuscripts that do not rigorously validate their methods generally will not be considered for publication in JCIM. Validation of methods should follow standard procedures with a clear and transparent training/test data splitting.

**Guo-Wei Wei** orcid.org/0000-0002-5781-2937
**Feng Zhu** orcid.org/0000-0001-8069-0053
**Kenneth M. Merz** orcid.org/0000-0001-9139-5893

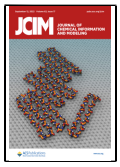
## AUTHOR INFORMATION

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jcim.2c01031

### Notes

Views expressed in this editorial are those of the authors and not necessarily the views of the ACS.

**Published:** August 31, 2022




© 2022 American Chemical Society

https://doi.org/10.1021/acs.jcim.2c01031
J. Chem. Inf. Model. 2022, 62, 3941−3941