DISCUSSION PUBLISHERS
Research Center · Open Access Edition

Machine learning methods in chemoinformatics

John B. O. Mitchell

Discussion Research ID	DPR-2026-1136
Year	2014
Field	Computer Science
DOI	10.1002/wcms.1183
Author ORCID	0000-0002-0379-6097
Licence	CC-BY
Original source	Wiley Interdisciplinary Reviews Computational Molecular Science
Permanent link	https://discussionpublishers.com/research/928ec74a-9970-4824-b53e-eb92524c06c1
Publisher	Discussion Publishers discussionpublishers.com

Republished by Discussion Publishers under the original CC-BY licence. Copyright remains with the original authors; please cite the original work.

Cite as: John B. O. Mitchell (2014). Machine learning methods in chemoinformatics. Discussion Publishers, DPR-2026-1136.

Abstract

Machine learning algorithms are generally developed in computer science or adjacent disciplines and find their way into chemical modeling by a process of diffusion. Though particular machine learning methods are popular in chemoinformatics and quantitative structure-activity relationships (QSAR), many others exist in the technical literature. This discussion is methods-based and focused on some algorithms that chemoinformatics researchers frequently use. It makes no claim to be exhaustive. We concentrate on methods for supervised learning, predicting the unknown property values of a test set of instances, usually molecules, based on the known values for a training set. Particularly relevant approaches include Artificial Neural Networks, Random Forest, Support Vector Machine, k-Nearest Neighbors and nave Bayes classifiers.

Keywords: machine learning, cheminformatics, QSAR, supervised learning, artificial neural networks, random for

Full text

The complete text of this work is hosted by its original publisher.
Read it via the Discussion Publishers page:
https://discussionpublishers.com/research/928ec74a-9970-4824-b53e-eb92524c06c1

Original record: https://doi.org/10.1002/wcms.1183
DOI: https://doi.org/10.1002/wcms.1183