Machine Learning the Structure and Reactivity of Transition Metal Complexes

Dr. David Balcells

Department of Chemistry, University of Oslo

Email: d.b.badia@kjemi.uio.no

This seminar will present an overview of the research done in my group in the area of machine learning, which focuses on the structure and reactivity of transition metal species. Our research approach is structured into three interacting blocks; namely 1) data, 2) representations, and 3) models. In the first block, we are working on the automated computation of datasets including the quantum properties of a large number of metal complexes. The tmQM dataset[1] represents our first effort in this direction, containing 86k complexes from the Cambridge Structural Database, for which we provided properties computed with semiempirical and DFT methods. Regarding representations, we are developing molecular graphs expressing electronic structure properties derived from NBO analysis.[2] Regarding models, we have worked on the prediction of energy barriers and transition state geometries that are relevant to catalysis.[3,4] In particular, we explored how deep, Bayesian, and ensemble methods can be combined for accurate, affordable, and explainable predictions, in a way that is useful for catalyst design.

References

[1] Balcells et al., J. Chem. Inf. Model., 2020, 60, 6135.
[2] Balcells et al., ChemRxiv preprint, 2022, 10.26434/chemrxiv-2022-fd43k-v2
[3] Aspuru-Guzik, Balcells et al., Chem. Sci., 2020, 11, 4584.
[4] Balcells, Ess et al., Top. Catal., 2021, 65, 312.