ETH zürich

Digital Chemistry FS2024

NMRCRAFT

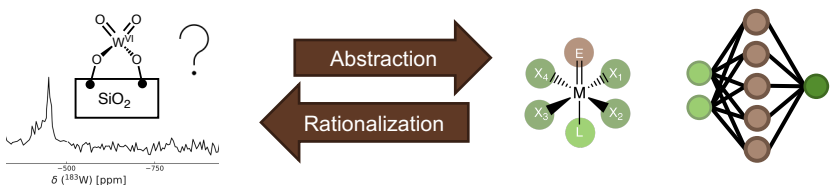
Crafting Catalysts from NMR Features

Karolina Biniek[1], Magdalena Lederbauer[1], Samuel Stricker[1], Yingnan Wang[1], Tiago Würthner[1]
[1]Department of Chemistry and Applied Biosciences, ETH Zurich, Vladimir-Prelog-Weg 1-5, 8093 Zurich

1 Introduction

Heterogeneous W and Mo catalysts drive olefin metathesis, an indispensable industrial process[1]

- **Problem:** Catalyst activation **mechanism** unknown[1]
- **Solution:** Spectroscopic tools (**NMR**) elucidate structure-reactivity relationships → Chemical shift tensors (**CSTs**) link to frontier molecular orbitals[2]

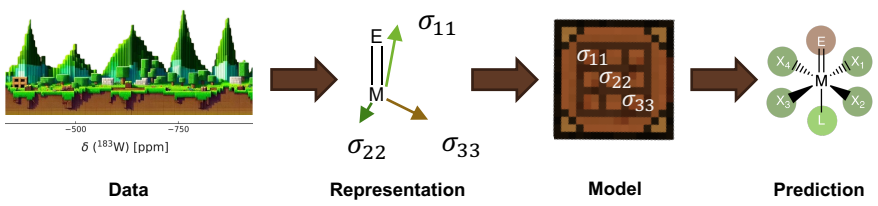


- **Challenges:** Limited **data**, unknown catalyst **structure**

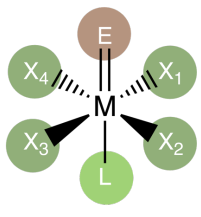
2 Methods

Previous work: Predicting NMR shifts from descriptors (i.e. properties)

This work: Predicting ligands from NMR shifts



- **Data:** 19’169 *in-silico* generated octahedral complexes and their DFT-computed CSTs



M = Mo, W

E = O, S, Se, NR

X_i = Cl, Br, OtBu, O^{t}Bu$_3$F, O^{t}Bu$_6$F, O^{t}Bu$_9$F, OCF$_3$, OSiF$_3$, OSi(O^{t}Bu)$_3$, OSiPh$_3$, OPh, OC$_6$F$_5$, O(2,6-Ph)Ph, SPh, Me, Et, ^{t}Bu, neopentyl

L = THF, OEt$_2$, pyridine

- **Representation:** NMR CSTs as numerical features, ligands as categorical features
- **Model Selection:** Classifiers[2]

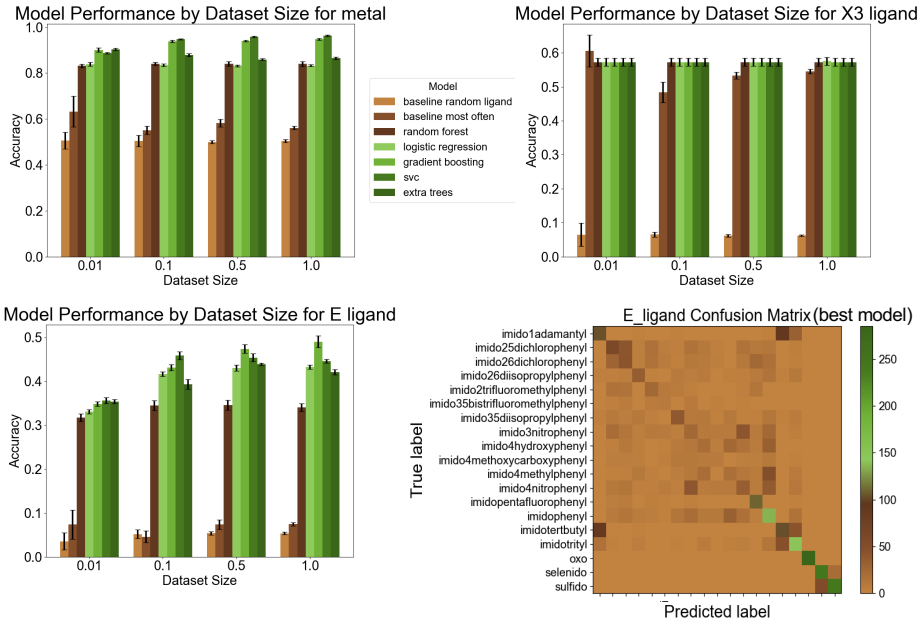
Gradient Boosting	Extra Trees
Logistic Regression	Random Forest
Support Vector	Dummy

[2]hyperparameter tuning using the package hyperopt[3]

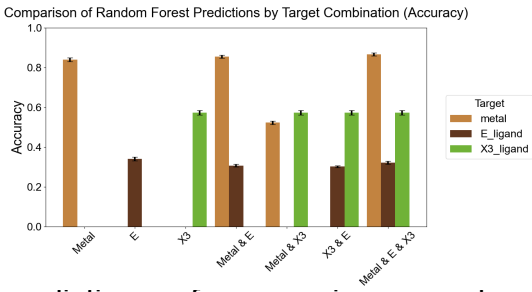
- **Evaluation:** Held-out test dataset using bootstrapping

3 Results and Discussion

- How well can a classifier predict the metal and single ligands from the NMR CSTs?



- Do the predictions improve when predicting multiple targets, i.e. the metal *and* a ligand instead of the metal or ligand *separately*?



- Does the predictive performance improve when incorporating other ligands as an additional input to the CSTs?

Target	With Ligands: Accuracy / %	Without Ligands: Accuracy / %
metal	91.4 ± 0.3	88.0 ± 0.5
E_ligand	43.5 ± 0.5	34.3 ± 0.8
X3_ligand	54.3 ± 1.1	57.3 ± 1.1

4 Conclusion

- First application of classifying ligands from NMR chemical shifts of a metal and a coordinating ligand atom (E), outperforming baselines
- More complex models (Neural Networks), data augmentation for X-ligands

References

[1] D. F. Nater, C. J. Kaul, L. Lätsch, H. Tsurugi, K. Mashima, C. Copéret, *Chemistry A European J.* 28 (**2022**).
[2] Z. J. Berkson, L. Lätsch, J. Hillenbrand, A. Fürstner, C. Copéret, *J. Am. Chem. Soc.* 144, 15020–15025 (**2022**).
[3] J. Bergstra, D. Yamins, D. Cox, in *Proceedings of the 30th International Conference on Machine Learning* (PMLR, 2013; https://proceedings.mlr.press/v28/bergstra13.html), pp. 115–123.