Data-Driven Discovery of Electrocatalysts for the CO2 Reduction Reaction: Getting Into the Right Shape

B. Ranković[1,3], L. E. Zaza[2,3], R. Buonsanti[2,3]*, P. Schwaller[1,3]*

[1]EPFL SB-ISIC LIAC, [2]EPFL SB ISIC LNCE, [3]NCCR Catalysis

In the face of ongoing climate change, harnessing carbon dioxide (CO2) and converting it to useful chemicals via electrocatalytic processes stands out as a promising strategy [1, 2]. However, a major challenge is the lack of efficient and selective catalysts. Copper, despite its suitable binding energy for generating a range of products, grapples with selectivity issues in its bulk form.

Our study tackles this challenge by utilizing machine learning and data-driven computational tools to accelerate catalyst discovery. Integrating Electronic Lab Notebooks (ELN) [3], we have devised a process that makes data sharing across different chemistry laboratories machine-learning readable. This collaborative approach enables us to catalogue and analyze important synthesis and catalysis parameters. Notably, we enhance the dataset covering our wet-lab experiments with curated literature data, ensuring a comprehensive dataset for models to learn from.

Our models demonstrate promising results, accurately predicting nanocrystal shapes from reaction vectors derived from reagents and reaction conditions, features significantly associated with electrocatalytic properties [4]. We can classify multiple and intermediate shapes using multi-output classification models. Prediction probabilities provide an overview of different shapes present in the reaction output, underscoring our ability to decipher the rules guiding the formation of specific shapes from reaction vectors.

This uncovered correlation between reaction vectors and resulting nanoparticle shapes validates our research direction towards Bayesian optimization strategies [5]. Bayesian search facilitates exploration of the reaction space for specific shapes known to enhance selectivity. Importantly, it is not limited to known conditions but can potentially guide reactions towards direct selectivity optimization. An intriguing byproduct of this approach could be the emergence of nanoparticle shapes that were not initially targeted but have proven beneficial for catalytic selectivity, emphasizing the adaptive potential of our methodology.

In summary, our work presents a progressive approach towards the data-driven discovery of electrocatalysts, with a focus on CO2 reduction. It promises significant strides not only in enhancing the efficiency and selectivity of catalysts but also in transforming the traditionally slow process of lab data analysis into a more dynamic, automated one, paving the way for faster and more accurate catalyst discovery.

[1] Jianfeng Huang, Raffaella Buonsanti, *Chemistry of Materials*, **2018**, 31.1: 13-25.
[2] Giulia Mangione, Jianfeng Huang, Raffaella Buonsanti, Clémence Corminboeuf, *The Journal of Physical Chemistry Letters*, **2019**, 10.15: 4259-4265.
[3] Raffaella Buonsanti, *Chemistry of Materials*, **2023**, 35.3: 805-806.
[4] Rachel M. Anderson, David F. Yancey, Liang Zhang, Samuel T. Chill, Graeme Henkelman, Richard M. Crooks, *Accounts of chemical research*, **2015**, 48.5: 1351-1357.
[5] Benjamin J. Shields, Jason Stevens, Jun Li, Marvin Parasram, Farhan Damani, Jesus I. Martinez Alvarado, Jacob M. Janey, Ryan P Adams, Abigail G. Doyle, *Nature*, **2021**, 590.7844: 89-96.