ChBE 6745 Project: Prediction of Adsorption Properties of Metal-Organic Frameworks with Framework Flexibility

Chao-Wen Chang, Pengfei Cheng, Po-Wei Huang, Xiaohan Yu, Yamin Zhang

Problem definition

Agrawal and Sholl (2019) show that appropriate consideration of framework flexibility may be important to quantitative predictions about molecular adsorption in metal-organic frameworks (MOFs). However, taking the framework flexibility into account directly in the molecular simulation framework may be 5-10 times morecomputationally expensive than standard simulation methods based on rigid crystal structure. Therefore, we are interested in constructing a machine learning model to predict the adsorption properties of MOFs with framework flexibility based on: (1) the features of the MOFs, (2) the features of adsorbants, and (3) the adsorption properties from standard simulation methods based on rigid crystal structure.

Dataset details

The dataset is in the `data` folder.
It is composed of three parts:

1. 36 MOF features (under `data/ML_data`)
2. 6 adsorbant features (manually added in `data_input.ipynb`)
3. adsorption uptakes of 882 (MOF, adsorbant) pairs (under `data/flexibility_data/y_data/adsorption_data`), containing two values:
   i. values from rigid model
   ii. mean values from flexible model

The data has been preliminary processed in `data_input.ipynb`.

Model training strategy

1. Multi-linear regression
2. Kernel regression
3. Kernel ridge regression
4. Lasso regression
5. Genetic programming
6. Neural network

Model validation strategy

1. hold-out
2. k-fold
3. bootstrapping

Reference

Agrawal, Mayank, and David S. Sholl. "Effects of Intrinsic Flexibility on Adsorption Properties of Metal–Organic Frameworks at Dilute and Nondilute Loadings." *ACS applied materials & interfaces* 11.34 (2019): 31060-31068.