Predicting reaction performance in sulfonamide electrochemical synthesis using machine learning

Yiran Cao & Timothy Noël

**Introduction**

Machine learning methods are becoming integral to scientific inquiry in numerous disciplines. It was demonstrated that machine learning can be used to predict the performance of a synthetic reaction in multidimensional chemical space using data obtained via experimentation.

**Project summary**

Using molecular descriptors as inputs and reaction yield as outputs, a random forest algorithm provides significantly improved predictive performance over linear regression analysis. The random forest model will also be applied to sparse training sets and out-of-sample prediction, suggesting its value in facilitating adoption of synthetic methodology.

**Project goals**

Optimize model pipeline using Python and ChemSAR platform

**Contact information**

Yiran Cao  y.cao@tue.nl  STO 1.39  Timothy Noël  t.noel@tue.nl  STO 1.37