

Title: Putting density functional theory to the test with machine learning
Author: Heather J. Kulik, Department of Chemical Engineering, MIT

Abstract: Accelerated simulation with machine learning (ML) has begun to provide the advances in efficiency to make property prediction tractable at an unprecedented scale. Nevertheless, ML-accelerated workflows both inherit the biases of training data derived from density functional theory (DFT) and leads to many attempted calculations that are doomed to fail. Many compelling molecular systems involve strained chemical bonds, open shell radicals and diradicals, or metal-organic bonds to open-shell transition-metal centers. Although promising targets, these materials present unique challenges for electronic structure methods and combinatorial challenges for their discovery. I will describe some of my group's recent advances in using artificial intelligence to address challenges in accuracy and efficiency beyond conventional DFT-based ML workflows. I will describe how we have developed ML models trained to predict the results of multiple methods or the differences between them, enabling quantitative sensitivity analysis. I will then describe ML models we have developed on a series of chemical and electronic structure descriptors that predict the likelihood of calculation success and detect the presence of strong correlation. Combining novel descriptors and developing consensus from multiple levels of theory empowers "decision engines" that represent the first steps toward autonomous workflows that avoid the need for expert determination of the robustness of DFT-based computational modeling.