Prospects for Bayesian optimization to reduce the complexity of materials processing optimization

Paulette Clancy
Dept. of Chemical and Biomolecular Engineering
The Johns Hopkins University
Baltimore MD 21218

Sept. 2020

There are many problems at the forefront of materials chemistry whose solution is stymied by its inherent complexity. Such problems are characterized by a rich landscape of parameters and processing variables that is combinatorially too large for either an experimental or a computational approach to solve through an exhaustive search. In such cases, the usual approach is an Edisonian trial-and-error approach, which inevitably leaves areas of parameter space unexplored. The problems that we have explored are also characterized by a scarcity of data, since the data are expensive to acquire, both experimentally and computationally. This makes it an ideal candidate to solve using a Bayesian optimization (BayesOpt) approach.

We have used a Bayesian approach to study several problems in materials chemistry. This talk will discuss the most mature test case in which we looked at “solvent engineering” to optimize the process by which some novel solar cells are fabricated. We used BayesOpt extensively to study how to optimize the choice of solvent or additive to produce high quality thin films of lead-based hybrid organic-inorganic perovskites. These “perovskites” all exhibit a crystalline habit with the formula ABX$_3$, where A and B are cations and X is an anion. Here, we chose B to be lead and X to be halides, and we allow A-site cations to be chosen among a small group of options. Using BayesOpt, we identified the best solvent out of a list of around 12 solvents, and also how to optimize the best mixed-halide system to solubilize the lead salt.

To validate these predictions, we used the Bayesian-generated ranked list of “best performers” to identify when a crystalline “intermediate” is likely to form, which can adversely affect thin film quality. We confirmed our ability to determine whether a precursor forms in six cases (three positive for crystalline intermediates and three negative for intermediates). To do so, we were fortunate to have XRD experiments performed by Lara Estroff’s group at Cornell that confirm our predictions.

I will end with some ideas of where the BayesOpt field can expand its use in the chemical sciences and share some “lessons learned” in implementing BayesOpt and machine learning, in general, in my group.