Advances in Bayesian Optimisation for materials design

Milica Todorovic[1]

[1]University of Turku, Finland

Materials research can be accelerated with active machine learning methods like Bayesian optimization (BO), where datasets are collected on-the-fly in the search for optimal outcomes. We encoded this probabilistic algorithm into the Bayesian Optimization Structure Search (BOSS) Python tool [1] and deployed it to study complex phenomena such as surface and interface chemistry, magnets, gas and biosensors, and in AI-guided experimental protocols. Active learning is now a well-established method in both computational and experimental research, achieving solutions previously inaccessible through human intuition.

However, materials acquisitions are costly and datasets small. Further algorithm development and transfer of knowledge from computer science is necessary to focus and accelerate active learning. In materials science, we demonstrate how data gradients and materials symmetry [2] can be leveraged to fit better models with less data. We used human-in-the-loop strategies to direct the search towards desired areas of parameter space [3]. Multi-task Gaussian Process models [4] were deployed to integrate information from different simulation types into multi-fidelity BO, so that less accurate data can inform the search at a higher level of accuracy. Domain-specific problems can drive algorithm development -- there are many challenges but also opportunities for cross-disciplinary innovation.

References:

[1] M. Todorović et al., npj Comput. Mater., 5, 35 (2019)
[2] J. Li et al., Small Struct. 2400268 (2024)
[3] A. Tiihonen et al.,, AI for Accelerated Materials Design NeurIPS 2022 Workshop (2022)
[4] E. V. Bonilla et al., Adv. Neural Inf. Process. 20 (2027)