** Materials Science and Engineering**

122 S. Central Campus Drive, Salt Lake City, Utah 84112 (801) 581-8632

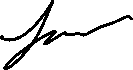
January 7, 2023

Dear Editor:

We are very pleased to submit this data article to Data in Brief. Our manuscript, titled *Materials Science Optimization Benchmark Dataset for High-dimensional, Multi-objective, Multi-fidelity Optimization of CrabNet Hyperparameters* provides important research from my research group at the University of Utah. This work presents a benchmark dataset for materials science optimization tasks that incorporates both heteroskedastic noise in a realistically complex setting with multiple objectives, multiple fidelities, constraints, and high dimensionality.

The dataset represents 387 RTX 2080-Ti days’ worth of GPU computation time (4614 CUDA core years) and contains over 150,000 datapoints. The dataset presented in this work is used to create a surrogate model as close as possible to running the actual training of the composition-based regression model. This will help form part of a larger suite of experimentally and computationally derived benchmarks. Additionally, this dataset can serve as an optimization task for advanced Bayesian optimization topics including multi-objective, multi-fidelity, high-dimensional, and linearly constrained optimization.

Sincerely,



Dr. Taylor Sparks  
Associate Professor & Associate Chair  
Materials Science and Engineering Department  
University of Utah  
Salt Lake City, Utah 84112