Abstract

We predicted the in-plane elastic stiffness of two-dimensional materials using only their chemical formula. We used the Computational 2D Materials Database (C2DB) to build a curated dataset of 7462 monolayers with averaged in-plane stiffness values and eight compositional features. We trained an XGBoost model and tested it fairly by ensuring that the same chemical formula did not appear in both the training and testing sets, and by using 3-fold cross-validation. The model explained most of the variation in $\bar{E}$, achieving ($R^2 = 0.749$) with relatively small errors (MAE = 20.9 N/m; RMSE = 33.4 N/m; $R^2_{\mathrm{CV}} = 0.767 \pm 0.007$). After that, we performed a guided search (Bayesian Optimization) in the standardized feature space to identify promising candidates. We accepted only points close to known materials: for each accepted point, we selected the closest neighboring real material in C2DB. By applying this distance constraint and simple chemical filtering, we obtained 52 materials ordered by their calculated DFT stiffness values in C2DB. Finally, we built a simple tool with a smooth workflow that triages potential materials with clear rules, fast predictions, and clear explanations, allowing prioritization of promising materials for further (more costly) experimental testing.