# DiSCoVeR: An Attention and Density-Based Machine Learning Algorithm for Discovering Novel Superhard Materials

Liu, Marianne (School: West High School)

This project presents the first automated screening method for predicting high-performance materials, with special emphasis on novel superhard materials. Superhard materials are key for many modern technology industries. Due to the high cost of diamond, diamond-like materials have long been sought after with limited results, however, high-throughput screening can accelerate this process. First, the hyperparameters of a machine learning model (CrabNet) were optimized, and a new state-of-the-art model within the materials informatics benchmarking platform was achieved with about a 4.5% decrease from the previous best performing model. Then, the DiSCoVeR algorithm was created by combining the state-of-the-art CrabNet model with three additional data analysis techniques: Element Mover's Distance, DensMAP, and HDBSCAN. DiSCoVeR enabled the screening of 70,000 compounds for novel superhard materials and three of the top candidates synthesized successfully experimentally validated DiSCoVeR's predictions. All three candidates exhibit significantly greater Vickers hardness than any existing hard materials used in industrial applications today. Finally, DiSCoVeR is employed to screen over 7 million unique candidates, which include completely theoretical compounds never synthesized before, several of which were predicted to have a bulk modulus over 100 GPa greater than diamond. DiSCoVeR is shown to successfully identify novel chemical compounds with desired material properties and also discover high-performing compounds within chemically similar compound classes. Thus, DiSCoVeR not only helps pave the way to identifying synthesizable materials that are harder and more cost-effective than diamond, but also demonstrates a method to accelerate materials discovery in other fields.

**Awards Won:**

Third Award of $1,000