Machine Learning for Prediction of Topological Materials for Chemical Sensing

Sepehr Roonasi
Department of Chemistry, Simon Fraser University, Burnaby BC Canada V5A 1S6

Haravneet Kaur
Department of Chemistry, Simon Fraser University, Burnaby BC Canada V5A 1S6

Paul C.H. Li
Department of Chemistry, Simon Fraser University, Burnaby BC Canada V5A 1S6

Abstract

The recently discovered topological insulator (TI) exhibits quantum properties that are highly sensitive and selective gas sensing without noise problems. This is because TI is insulating in its bulk but conducting uni-directionally on its boundaries. The materials properties of known TIs, such as Bi2Se3, will be used to train the convoluted neural network (CNN) used in machine learning (ML). In order to optimize sensor performance, new TI materials will be predicted using ML.

ML methods in artificial intelligence (AI) are becoming increasingly popular in accelerating the design of new materials by predicting material properties with accuracy close to ab initio calculations, but with computational speeds of orders of magnitude faster. An orbital graph convolutional neural network (OGCNN) has been adapted to directly learn material properties from the connection of atoms in the crystal. This method has provided a highly accurate prediction of DFT-calculated properties for different properties of crystals with various structure types and compositions after being trained with data obtained from Materials Project database. The material properties include formation energy, band gap, Fermi energy, spillage factor, etc.

The material structure is represented by a crystal graph, in which the nodes represent atoms, and the edges represent connections between atoms in a crystal. The atom (node) features include group and period numbers, valence electrons, ionization energy, electron affinity, atomic volume; the bond (edge) feature includes atom distance. In addition to crystal graph, orbital-orbital interactions are also included by the two-dimensional descriptor called the orbital field matrix (OFM). The inclusion of these interactions to encode the local chemical environments of atoms along with embedding of an encoder-decoder network enabled the OGCNN to achieve higher accuracy for prediction of new topological sensing materials.

Full details: Recently discovered topological insulator (TI), such as Bi2Se3, exhibits quantum properties that are highly sensitive and selective gas sensing without noise problems. This is because Bi2Se3 is insulating in its bulk but conducting uni-directionally on its boundaries. The composition of known TIs, such as Bi2Se3, will be used to train the convoluted neural network (CNN) used in machine learning (ML). In order to optimize sensor performance, new TI materials will be predicted using ML.

ML methods in artificial intelligence (AI) are becoming increasingly popular in accelerating the design of new materials by predicting material properties with accuracy close to ab initio calculations, such as density functional theory (DFT), but with computational speeds of orders of magnitude faster. An orbital graph convolutional neural network (OGCNN) has been adapted to directly learn material properties from the connection of atoms in the crystal, providing a universal and interpretable representation of crystalline materials. This method has provided a highly accurate prediction of DFT-calculated properties for different properties of inorganic crystals with various structure types and compositions after being trained with 104 data points obtained from Materials Project database. For instance, the inorganic crystal structure database has materials covering 87 elements, 7 lattice systems, and 216 space groups. The material properties include formation energy, band gap, Fermi energy, bulk & shear moduli, Poisson ratio, spillage factor, etc.

The material structure is represented by a crystal graph, in which the nodes represent atoms, and the edges represent connections between atoms in a crystal. The atom (node) features include group and period numbers, electronegativity (0.5-4.0), covalent radius 25-250 pm), valence electrons (1-12), first ionization energy (1.3-3.3 eV), electron affinity (-3 to -3.7 eV), atomic volume (1.5-4.3 cc/mol); the bond (edge) feature includes atom distance (0.7-5.2 Å). In addition to crystal graph, orbital-orbital interactions are also included by the two-dimensional descriptor called the orbital field matrix (OFM). Each dimension consists of 32 orbital fields (s1-2, p1-6, d1-10, f1-14). The inclusion of these interactions to encode the local chemical environments of atoms along with embedding of an encoder-decoder network enabled the OGCNN to achieve higher accuracy for prediction of new topological sensing materials.