Searching Design Principles using Machine Learning from Physics to Materials Sciences

Sooran Kim[1]

[1]Kyungpook National University, Daegu, South Korea

Recently, data-driven machine learning (ML) has been used to predict material properties without requiring detailed mechanisms. While effective, ML often lacks interpretability. This can be improved through ML-guided empirical formulas or feature importance analysis, so-called interpretable ML. In this talk, we introduce interpretable ML models for two systems: cuprates and solid electrolytes. Firstly, we predict the maximum superconducting transition temperature ($T_{c,max}$) of hole-doped cuprates with the root-mean-square-error (RMSE) of 3.705 K and R^2 of 0.969. We propose an ML-guided empirical formula for $T_{c,max}$ using key descriptors such as Bader charge of apical oxygen, the bond strength between apical atoms, and the number of superconducting layers. Furthermore, we predict the $T_{c,max}$ of hypothetical cuprates generated by replacing apical cations with other elements. Secondly, we meticulously curated features considering dynamic properties and developed ML models to predict the ionic conductivity of solid electrolytes. We compiled phonon-related descriptors along with descriptors related to the structure and electronic properties. Our logistic regression classifiers exhibit an accuracy of 93%, while the random forest regression model yields a R^2 of 0.710. From feature analysis, we suggest that phonon-related features are essential for estimating the ionic conductivities in both models. Furthermore, we applied our prediction model to screen 264 Li-containing materials and identified 11 promising candidates as potential superionic conductors.