Contribution ID: **20** Type: **Oral talk**

Predicting superconductors critical temperature via machine leaning methods

Thursday, 12 June 2025 16:20 (20 minutes)

The intriguing phenomenon of superconductivity has been extensively studied for over a century. Yet, a key challenge remains unresolved in practice: understanding and predicting the critical temperature T_c of a superconductor. This issue is especially challenging for high-temperature superconductors (HTSC), which comprise diverse material classes and probably involve different electron pairing mechanisms.

In this work, we apply machine learning to address this challenge using the 3DSC dataset [1], with T_c values sourced from the SuperCon database [2]. The dataset integrates several descriptor types: the DSOAP representation – a modified Smooth Overlap of Atomic Positions (SOAP) framework that enables modeling of doped and non-stoichiometric compounds, which are especially common among HTSCs; basic atomic attributes from the MAGPIE descriptor set [3]; and materials properties computed via DFT from the Materials Project (MP) [4].

We explore dimensionality reduction via Principal Component Analysis (PCA), unsupervised clustering of superconductors, and optimization of three regression models – K-nearest neighbors, Random Forest, and Gradient Boosting – to predict T_c. Despite the limited data (less than 4000 compounds, including around 900 HTSCs), our models achieve competitive R^2 values, with the best scores exceeding 0.89.

A novel aspect of our approach is the integration of electronic structure information. We introduce features based on the electronic density of states (DOS), derived from a reworked DOS fingerprint [5], as well as a simple descriptor counting the number of bands crossing the Fermi level. Both sets of features are obtained from the MP. The inclusion of DOS-based features and information about band-crossing provide a richer representation of electronic structure, which is often overlooked. This integrated approach has the potential to support large-scale screening of DFT-computed materials databases, helping to identify promising candidates for new superconductors.

References

[1] T. Sommer, R. Willa, J. Schmalian, and P. Friederich, "3DSC - a dataset of superconductors including crystal structures," Sci Data, vol. 10, no. 1, Nov. 2023, doi: 10.1038/s41597-023-02721-y.
[2] V. Stanev et al., "Machine learning modeling of superconducting critical temperature," npj Comput Mater, vol. 4, no. 1, Jun. 2018, doi: 10.1038/s41524-018-0085-8.
[3] L. Ward, A. Agrawal, A. Choudhary, and C. Wolverton, "A general-purpose machine learning framework for predicting properties of inorganic materials," npj Comput Mater, vol. 2, no. 1, Aug. 2016, doi: 10.1038/npjcompumats.2016.28.
[4] A. Jain et al., "Commentary: The Materials Project: A materials genome approach to accelerating materials innovation," APL Materials, vol. 1, no. 1, Jul. 2013, doi: 10.1063/1.4812323.
[5] M. Kuban, S. Rigamonti, M. Scheidgen, and C. Draxl, "Density-of-states similarity descriptor for unsupervised learning from materials data," Sci Data, vol. 9, no. 1, Oct. 2022, doi: 10.1038/s41597-022-01754-z.

Primary authors: Mr SVYNARSKYI, Ihor (Kyiv Academic University); Dr BEZGUBA, Volodymyr (Kyiv Academic University)

Presenter: Mr SVYNARSKYI, Ihor (Kyiv Academic University)

Session Classification: Artificial Intelligence in Physics

Track Classification: Artificial Intelligence in Physics