Science Advances
AAAS

Supplementary Materials for

Machine learning of material properties: Predictive and interpretable multilinear models

Alice E. A. Allen* and Alexandre Tkatchenko*

*Corresponding author. Email: aliceeaallen@gmail.com (A.E.A.A.); alexandre.tkatchenko@uni.lu (A.T.)

Published 6 May 2022, *Sci. Adv.* **8**, eabm7185 (2022)
DOI: 10.1126/sciadv.abm7185

This PDF file includes:

Figs. S1 to S5

Fig. S1. The convergence of the H-statistics.

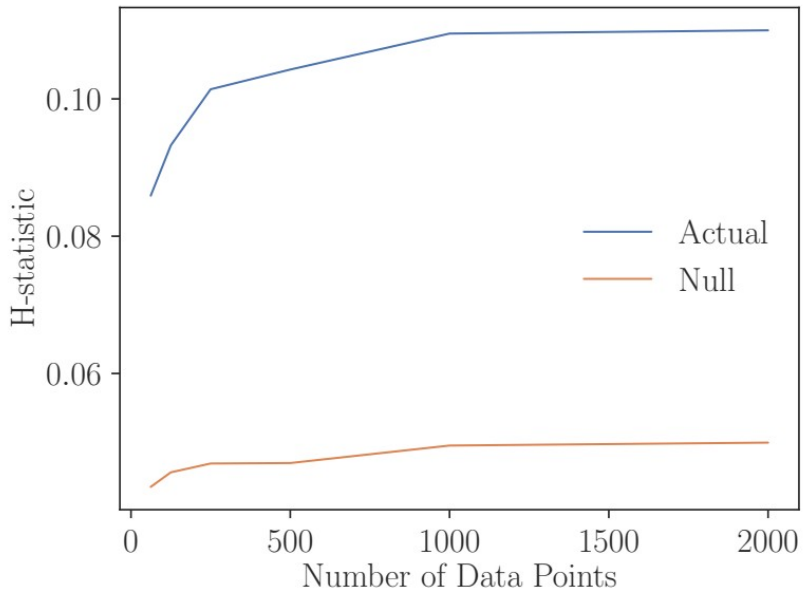


Figure S1: The convergence of the H-statistic with the number of data-points for the KRR model for the formation energy.

Fig. S2. The 2-gram interactions for the KRR model for TCOs.

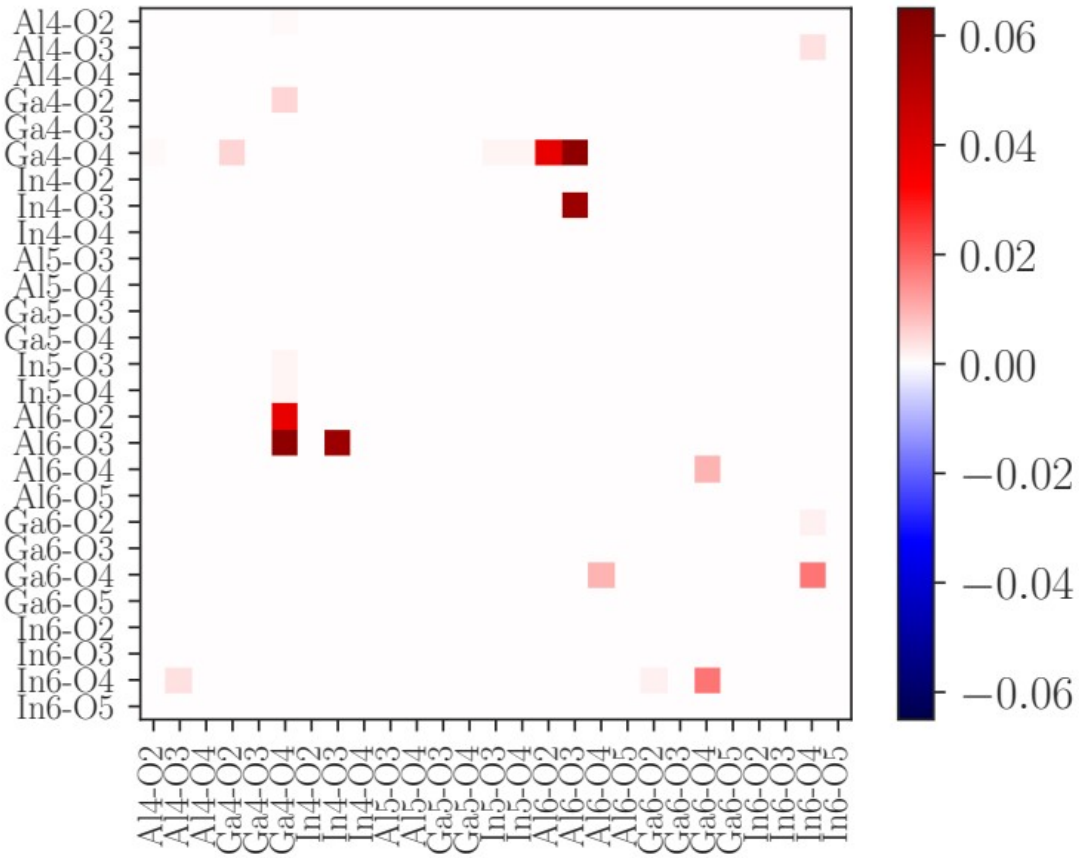
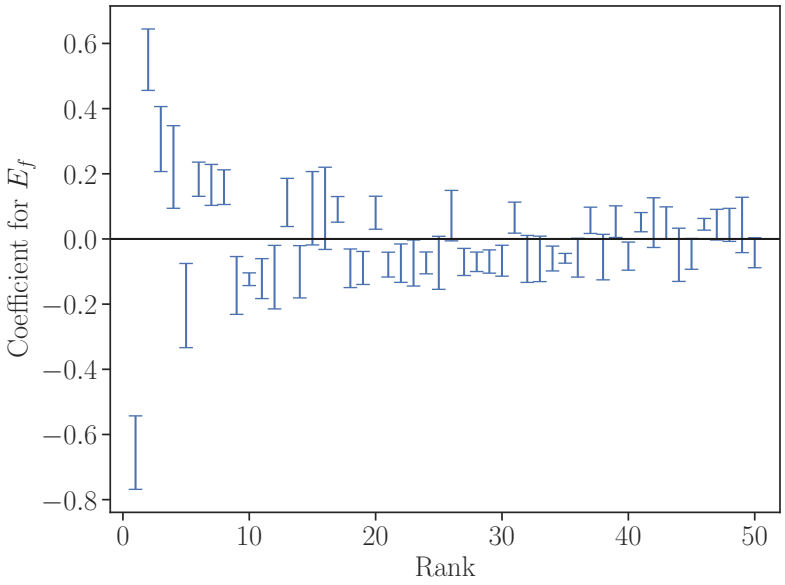


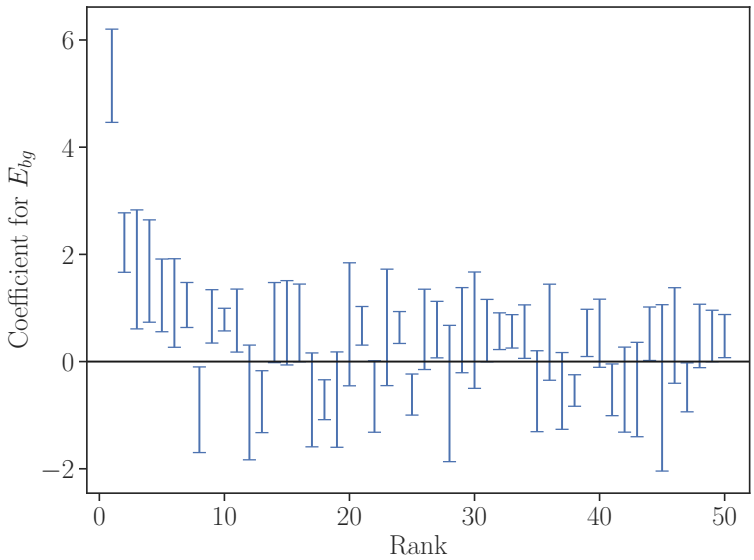
Figure S2: The interactions present for the 2-gram features for the 4-gram formation energy kernel ridge regression model.

Fig. S3. Analysis of the coefficients of the TCOs linear model. Figure S3: The top 50 coefficients for prediction of the a) formation energy and b) bandgap for TCOs. The variation across an ensemble of 10 linear models, fit from bootstrapped data, is shown. Part c) shows the coefficients that are both in the top 50 formation energies and top 50 bandgap energies and part d) shows the relationship between the concentration of O5 and the bandgap energy.

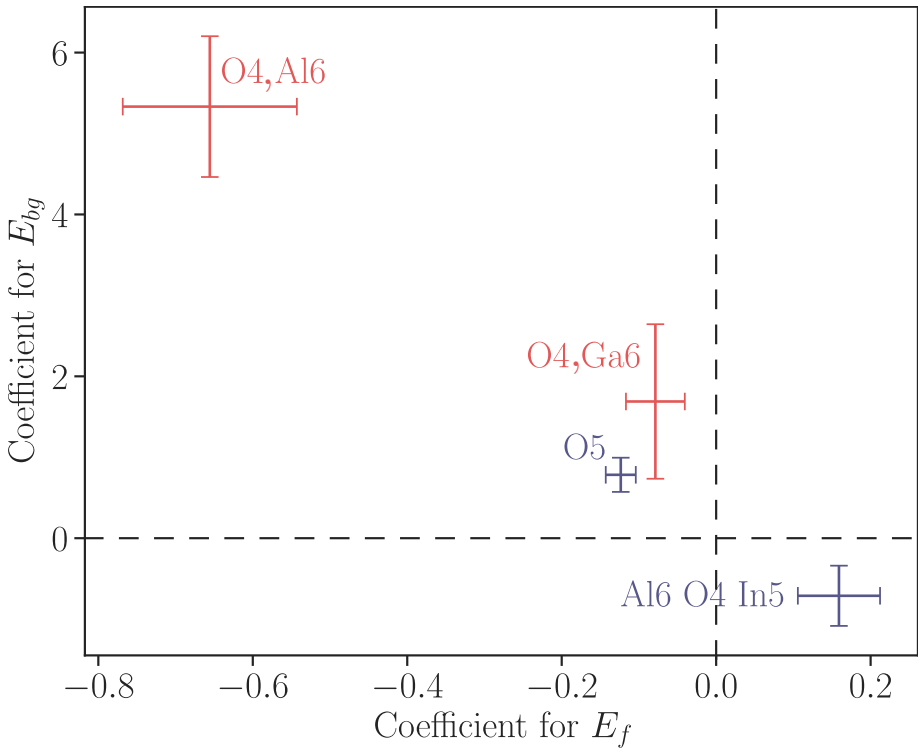
a)



b)



c)


O4,Al6
O4,Ga6
O5
Al6 O4 In5
Coefficient for E_{bg}
Coefficient for E_f
6
4
2
0
−0.8
−0.6
−0.4
−0.2
0.0
0.2

d

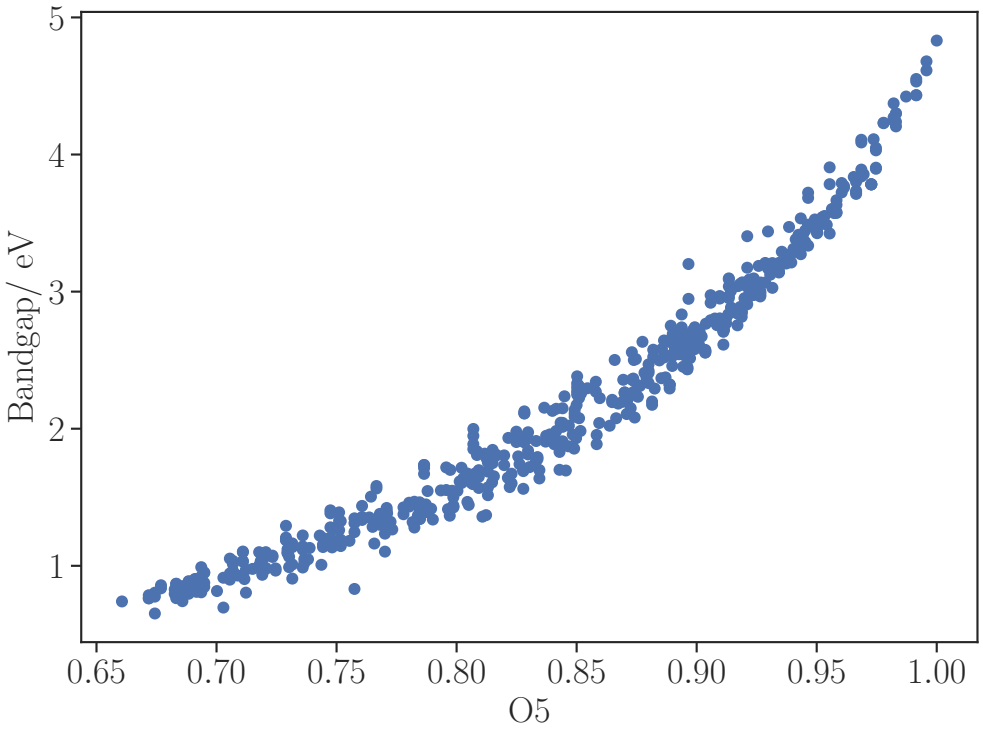

Bandgap/ eV
O5
5
4
3
2
1
0.65
0.70
0.75
0.80
0.85
0.90
0.95
1.00

Fig. S4. The coefficients for the elpasolite linear model.

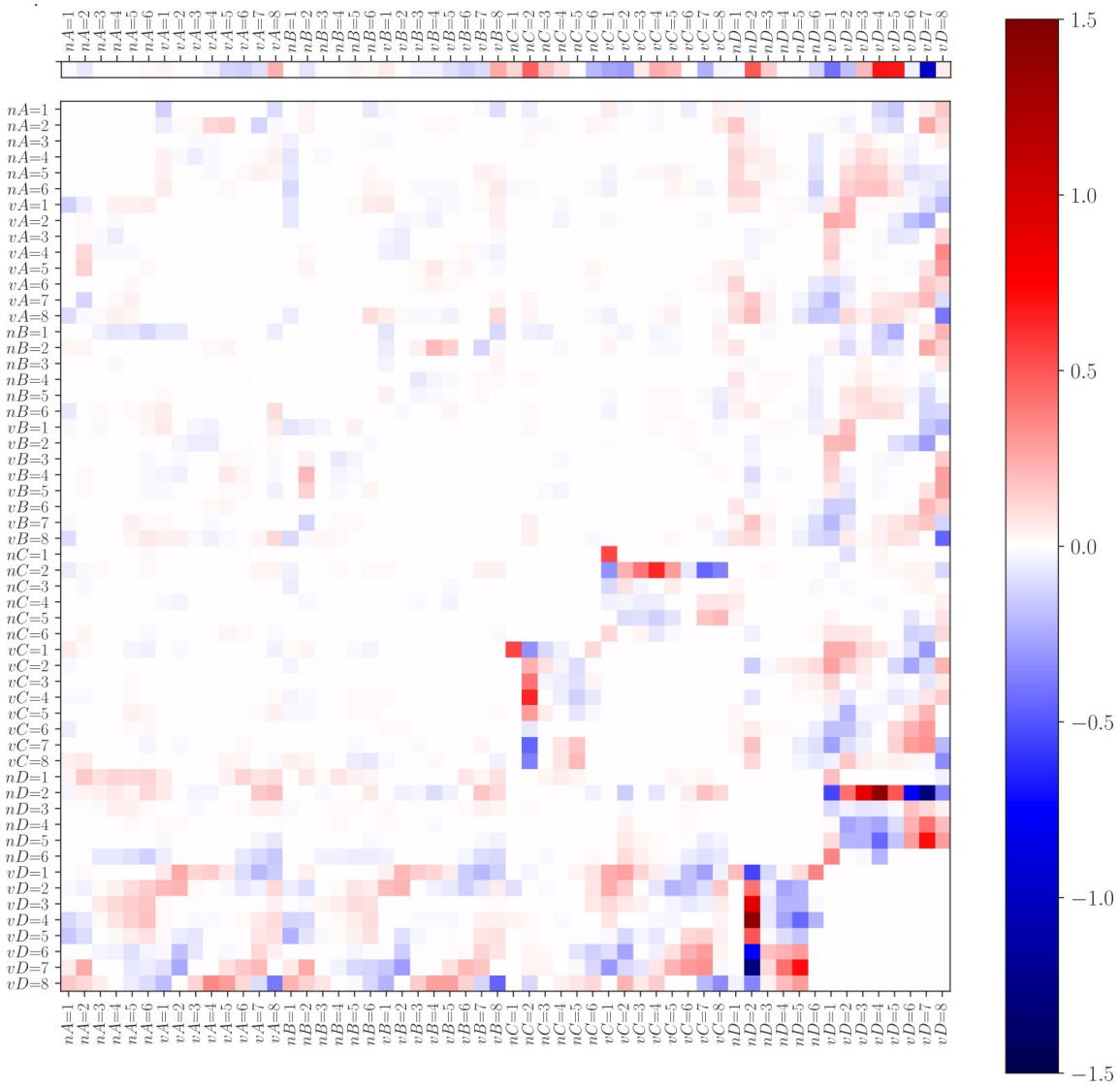


Figure S4.The coefficients for the individual variables and pairwise interaction terms for the trilinear model.

Fig. S5. Algorithms for exploring the elpasolite universe.

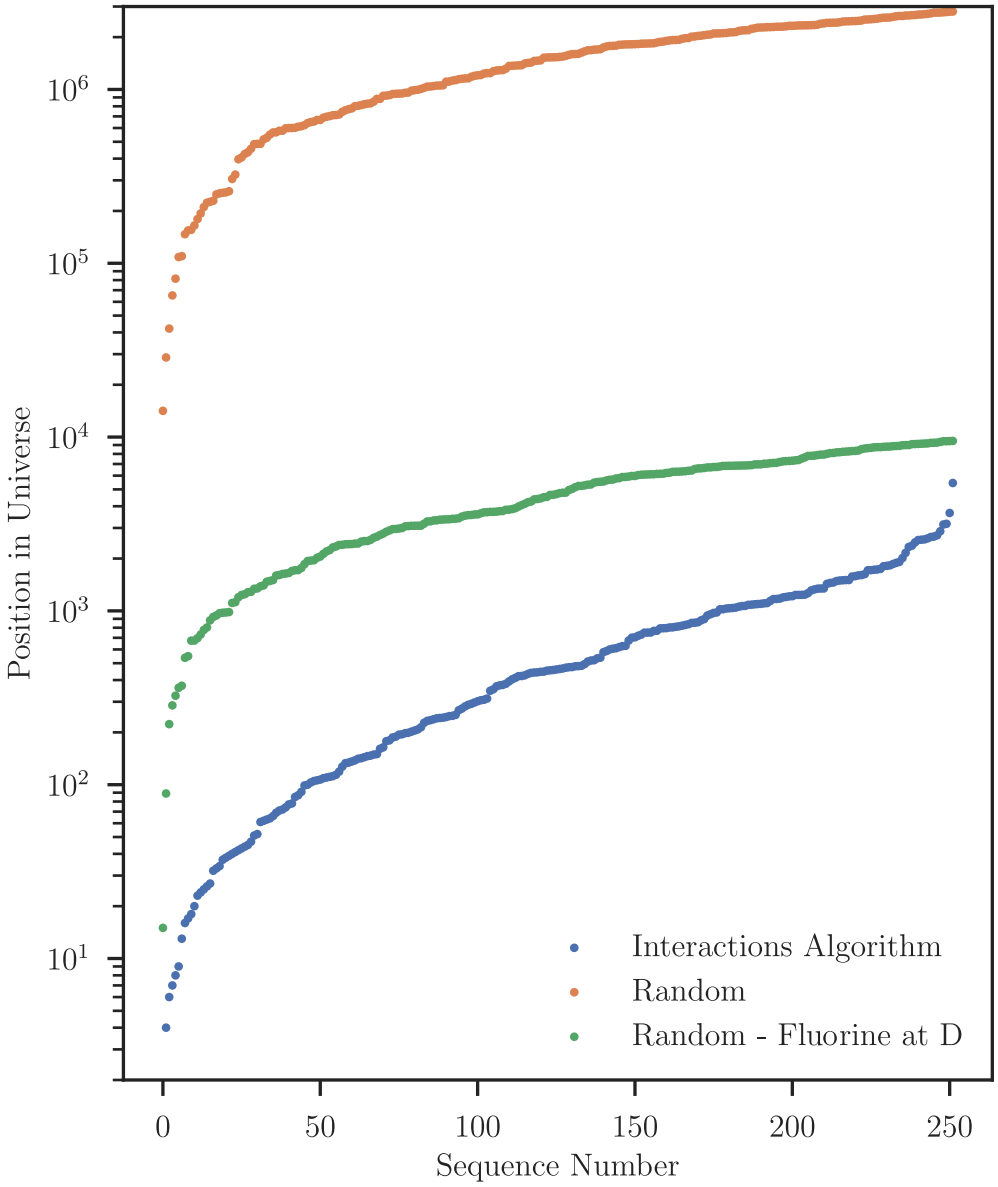


Figure S5.The position in the elpasolite universe of sequences from three different methods. The interaction algorithm uses three variable contributions to find low energy structures. The positions if the structures are randomly chosen are shown as are the positions if the structures are randomly chosen but restricted to having fluorine at D.