Supplementary Information: Discovering Novel Lead-Free Mixed Cation Hybrid Halide Perovskites via Machine Learning

Fatemeh Jamalinabijan,[†] Somayyeh Alidoust,[†] Gözde İniş Demir,[†] and Adem Tekin[*,†,‡]

†*Informatics Institute, Istanbul Technical University, 34469 Maslak, Istanbul, Türkiye*

‡*TÜBİTAK Research Institute for Fundamental Sciences, 41470 Gebze, Kocaeli, Türkiye*

E-mail: adem.tekin@itu.edu.tr

List of Figures

List of Tables

Table S1: Names and symbols of the features used in ML model generations for cubic, orthorhombic, and tetragonal phases

Feature Name	Feature Symbol	$\text{Model}^{E_{dec}}_{cubic}$	$\text{Model}^{E_{gap}}_{cubic}$	$\text{Model}^{E_{dec}}_{ortho}$	$\text{Model}^{E_{gap}}_{ortho}$	$\text{Model}^{E_{dec}}_{tetra}$	$\text{Model}^{E_{gap}}_{tetra}$
Electronegativity of B	χ_B	+	+	+	+	+	+
Electronegativity of X	χ_X	+	+	+	+	+	+
Radius of A	r_A	+	+	+	+	+	+
Radius of A'	$r_{A'}$	+	+	+	+	+	+
Radius of B	r_B	+	+	+	+	+	+
Standard deviation of thermal conductivity	$\sigma(\kappa)$	–	+	+	+	+	+
Standard deviation of mendeleev number	$\sigma(MN)$	+	+	+	+	–	+
Mean of mendeleev number	$\overline{MN}$	–	–	–	–	+	–
Standard deviation of X	$\sigma(X)$	+	+	+	+	–	+
Mean of X	$\overline{X}$	–	–	–	–	+	–
Standard deviation of atomic mass	$\sigma(m)$	–	+	+	+	+	+
Standard deviation of melting point	$\sigma(T_m)$	+	+	+	+	–	+
Mean of melting point	$\overline{T_m}$	–	–	–	–	+	–
Tolerance factor	t	+	+	+	+	+	+
New tolerance factor	T	+	+	+	+	+	+
Octahedral factor	μ	+	+	+	+	+	+
Mean of block	$\overline{b}$	–	+	–	+	–	+
Standard deviation of block	$\sigma(b)$	+	–	+	–	+	–
Standard deviation of atomic radius	$\sigma(r)$	+	–	–	–	–	–
Mean of group	$\overline{g}$	+	–	+	–	–	–
Standard deviation of group	$\sigma(g)$	–	–	–	–	+	–

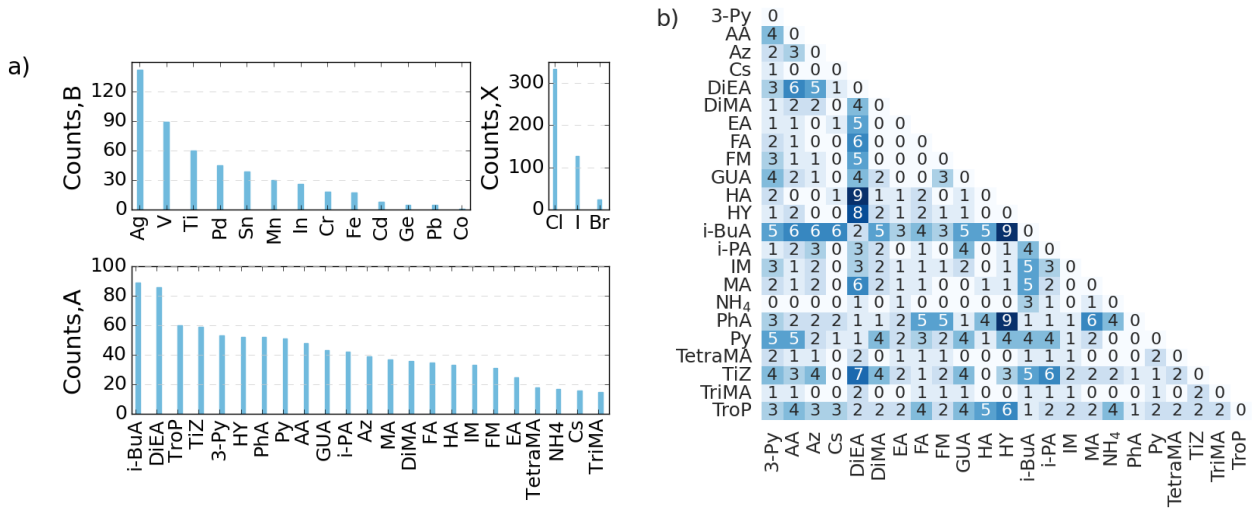

Figure S1. Distribution and frequency of occurrences of the A, B, and X components (left) and a heatmap analysis showing the distribution of A/A'-cations (right) within the 485 compounds.

Table S2: Errors of the E_{dec} and E_{gap} ML models for orthorhombic and tetragonal phases

Model Name	MSE_{train}	MSE_{test}	$\text{R}^2{}_{train}$	$\text{R}^2{}_{test}$	RMSD_{train}	RMSD_{test}	MAD_{train}	MAD_{test}
$\text{Model}^{E_{dec}}_{ortho}$	0.266	0.269	0.902	0.899	0.516	0.519	0.198	0.288
$\text{Model}^{E_{gap}}_{ortho}$	0.105	0.182	0.973	0.944	0.324	0.427	0.180	0.313
$\text{Model}^{E_{dec}}_{tetra}$	0.255	0.282	0.904	0.879	0.505	0.531	0.195	0.319
$\text{Model}^{E_{gap}}_{tetra}$	0.080	0.193	0.977	0.946	0.282	0.439	0.150	0.298

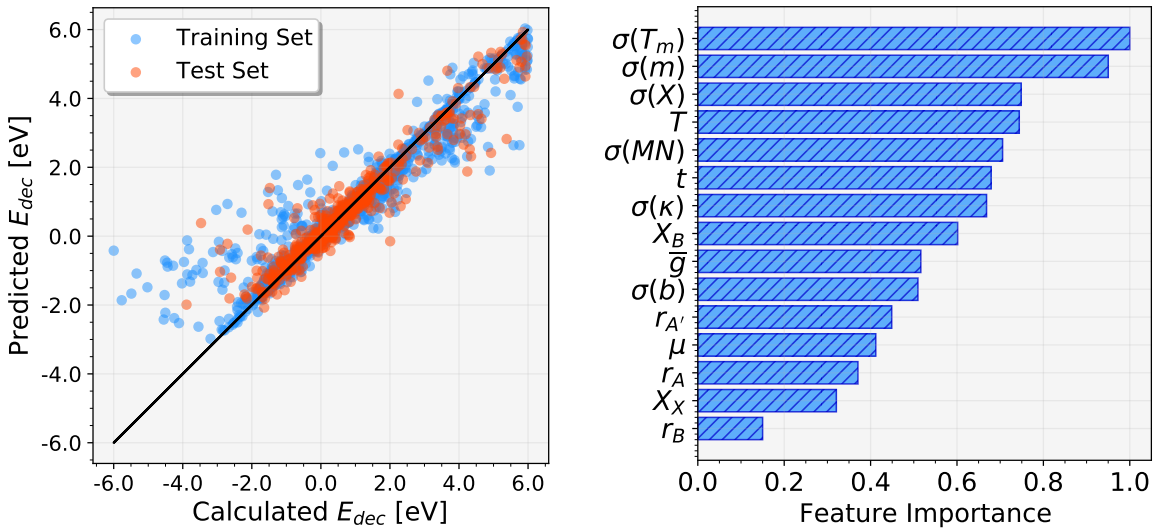

Figure S2. ML predicted decomposition energies of the orthorhombic perovskites (left) and the corresponding feature importance (right).

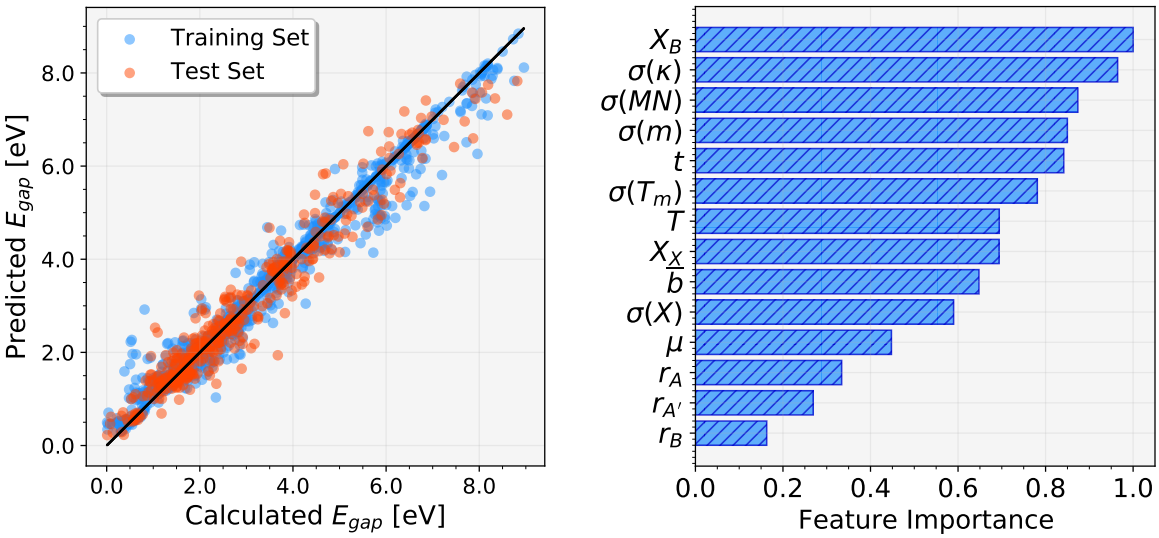

Figure S3. ML predicted band gap energies of the orthorhombic perovskites (left) and the corresponding feature importance (right).

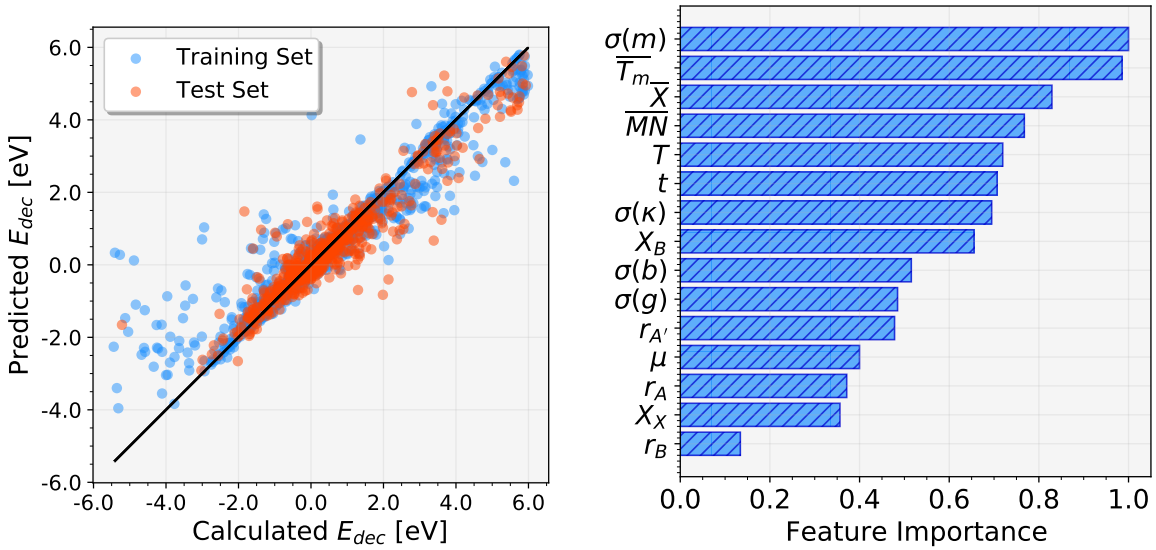

Figure S4. ML predicted decomposition energies of the tetragonal perovskites (left) and the corresponding feature importance (right).

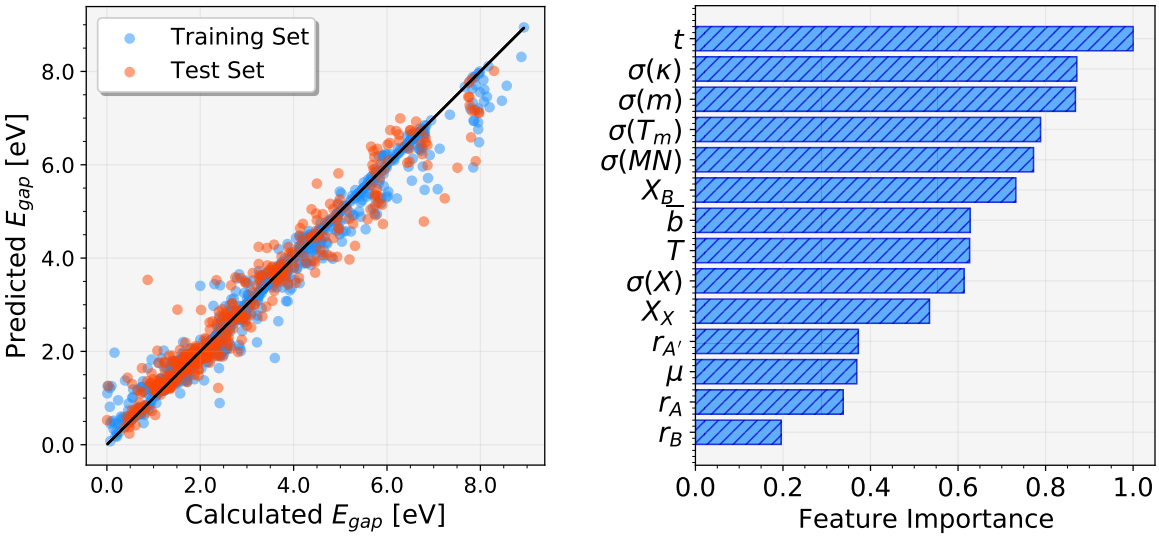

Figure S5. ML predicted band gap energies of the tetragonal perovskites (left) and the corresponding feature importance (right).