Electronic Supplementary Material (ESI) for Materials Horizons.
This journal is © The Royal Society of Chemistry 2023

Supplementary Information

Distilling Universal Activity Descriptors for Perovskite Catalysts from Multiple Data Sources via Multi-task Symbolic Regression

Zhilong Song[a,1], Xiao Wang[c,1], Fangting Liu[a], Qionghua Zhou[a,d]*, Wan-Jian Yin[b,e]*, Hao Wu[c], Weiqiao Deng[c]*, and Jinlan Wang[a,d]*

[a]School of Physics, Southeast University, Nanjing 211189, China.

[b]College of Energy, Soochow Institute for Energy and Materials InnovationS (SIEMIS), Soochow University, Suzhou 215006, China.

[c]Institute of Molecular Sciences and Engineering, Institute of Frontier and Interdisciplinary Science, Shandong University, Qingdao, Shandong 266237, China.

[d]Suzhou Laboratory, Suzhou, China

[e]Light Industry Institute of Electrochemical Power Source, Soochow University, Suzhou 215006, China.

*Corresponding authors.

E-mail addresses: qh.zhou@seu.edu.cn (Q. Zhou), wjyin@suda.edu.cn (W. Yin), dengwq@sdu.edu.cn (Q. Deng), jlwang@seu.edu.cn (J. Wang).

[1]These authors contributed equally: Zhilong Song, Xiao Wang.

Supplementary Contents

Supplementary Notes

Note S1 Original result of implementing MT-SR on 12 datasets

The best one-dimensional (1D, eq. 1) and two-dimensional (2D, eq. 2) descriptors obtained by MT-SISSO were (a 2D descriptor is a linear combination of two 1D descriptors),

$$\frac{R_{\mathrm{B}} + B_{\mathrm{AR}}}{t + A_{\mathrm{EG}}} \tag{1}$$

$$a\frac{N_{\mathrm{d}} R_{\mathrm{A}}}{Q_{\mathrm{A}} - R_{\mathrm{A}}} + b||N_{\mathrm{d}} - \chi_{\mathrm{B}}| - |N_{\mathrm{d}} - A_{\mathrm{FI}}|| \tag{2}$$

where *a* and *b* are constants. The average PCC of 1D and 2D descriptors on 12 datasets were shown in Table S2. The overall performances of the best 1D and 2D descriptors are shown in Fig. S1 and Fig. S2, respectively. The 2D descriptor showed a better linear relationship with OER activity than the 1D descriptor but still lacked accuracy on some datasets. The performances of these two descriptors on the independent testing set are shown in Fig. S3.

Note S2 Supplementary result of implementing MT-SR with *k*-fold CV

The best 1D (eq. 3) and 2D (eq. 4) descriptors produced by the 3-fold cross-validation method on 12 datasets:

$$\frac{A_{\mathrm{MP}} - N_{\mathrm{d}}}{e^{Q_{\mathrm{A}}}} \tag{3}$$

$$a|\frac{N_{\mathrm{d}}}{Q_{\mathrm{A}}} - \frac{A_{\mathrm{MP}}}{R_{\mathrm{A}}}| + b\frac{A_{\mathrm{MP}} A_{\mathrm{FI}}}{|\chi_{\mathrm{B}} - R_{\mathrm{AR}}|} \tag{4}$$

The best 1D (eq. 3) descriptors produced by the 3-fold cross-validation method on 16 datasets.

$$\frac{N_{\mathrm{d}}}{A_{\mathrm{FI}} e^{Q_{\mathrm{A}}}} \tag{5}$$

Note S3 Bayesian optimization

The Bayesian optimization python package[1] was used to optimize the parameters in descriptor 6, the main hyper-parameters setting and their meanings are shown in Table S4.

Note S4 Calculation of ranking accuracy and rank-biased overlap

For two ranking lists A and B, the ranking accuracy is calculated by:

$$RA = \frac{\sum_{i,j}^{k} \delta_{A_i,B_j}}{k^2-k},\ \delta_{A_i,B_j} = \begin{cases} 1, i \leqslant j \ and \ A_i \leqslant B_j \\ 1, i > j \ and \ A_i > B_j \\ 0, \ else \end{cases} \tag{6}$$

where k is the length of lists A and B, A_i and B_j is the i-th and j-th element in lists A and B, respectively.

The rank-biased overlap measures the similarity between two ranking lists[2]:

$$RBO = (1-p)\sum_{d=1}^{k} p^{d-1} \frac{|A_{1:d} \cap B_{1:d}|}{d} \tag{7}$$

where d is the depth of lists A and B, p is an adjustable parameter and the $A_{1:d}$ and $B_{1:d}$ means the 1:d-th element in lists A and B, respectively. $|A_{1:d} \cap B_{1:d}|$ is the proportion of common elements between two rankings lists $A_{1:d}$ and $B_{1:d}$.

Note S5 Calculation of overpotential

In order to consider the influence of temperature, the Gibbs free energy difference ΔG of O, OH and OOH before and after adsorption on the surface was calculated, according to eqs. 5-7:

$$\begin{aligned} \Delta G_{O^*} &= \Delta G(H_2O(g) + * \ \rightarrow O^* + H_2(g)) \\ &= G_{O^*} + G_{H_2} - G_{H_2O} - G_* \end{aligned} \tag{8}$$

$$\begin{aligned} \Delta G_{OH^*} &= \Delta G(H_2O(g) + * \ \rightarrow OH^* + 1/2H_2(g)) \\ &= G_{OH^*} + 0.5G_{H_2} - G_{H_2O} - G_* \end{aligned} \tag{9}$$

$$\begin{aligned} \Delta G_{OOH^*} &= \Delta G(2H_2O(g) + * \ \rightarrow OOH^* + 3/2H_2(g)) \\ &= G_{OOH^*} + 1.5G_{H_2} - 2G_{H_2O} - G_* \end{aligned}$$

(10)

where $G*$, G_{O*}, G_{OH*}, G_{OOH*}, G_{H_2} and G_{H_2O} are the free energy of the surface, the surface after O, OH, OOH absorption, H_2 and H_2O, respectively. The standard hydrogen electrode approximation was adopted to avoid calculating the energy of electrons. The free energy was calculated by eq. 8:

$$G = E + ZPE - TS$$

(11)

where E is energy, ZPE and S are the zero-vibration energy and entropy, respectively, which can be obtained by VASP vibrational calculation. T is the temperature setting to 300K. The free energies of O*, OH* and OOH* are replaced by energy E.

Thus, the free energy differences between the four reaction paths can be obtained:

$$\Delta G_1 = \Delta G_{\mathrm{OH}^*} \tag{12}$$

$$\Delta G_2 = \Delta G_{\mathrm{O}^*} - \Delta G_{\mathrm{OH}^*} \tag{13}$$

$$\Delta G_3 = \Delta G_{\mathrm{OOH}^*} - \Delta G_{\mathrm{O}^*} \tag{14}$$

$$\Delta G_4 = 4.92 - \Delta G_{\mathrm{OOH}^*} \tag{15}$$

Finally, the overpotential was calculated according to the free energy difference,:

$$\eta = max(\Delta G_1, \Delta G_2, \Delta G_3, \Delta G_4)/e - 1.23\,\mathrm{eV} \tag{16}$$

the free energy difference before and after OER is 4.92eV, so the ideal free energy difference of each step is 4.92/4 = 1.23eV.

Supplementary Tables

Table S1. 33 features used in this work and their abbreviations and meanings.

Feature	Meaning	Feature	Meaning	Feature	Meaning
R_A	A site ionic radius	A_{BP}	A site boiling point	B_{VRU}	B site UFF radius
R_B	B site ionic radius	B_{BP}	B site boiling point	A_{VRM}	A site MM3 radius
t	Tolerance factor	A_{DE}	A site density	B_{VRM}	B site MM3 radius
N_d	d electron number	B_{DE}	B site density	A_{EG}	A site ghosh electronegativity
χ_A	A site electronegativity	A_{LC}	A site lattice constant	B_{EG}	B site ghosh electronegativity
Q_A	A site valence	B_{LC}	B site lattice constant	A_{VRA}	A si7lvarezrez radius
χ_B	B site electronegativity	A_{MT}	A site melting point	B_{VRA}	B si7lvarezrez radius
A_{AW}	A site atomic mass	B_{MT}	B site melting point	A_{FI}	A site first ionization energy
B_{AW}	B site atomic mass	A_{VR}	A site Van der Waals radius	B_{FI}	B site first ionization energy
A_{AR}	A site atomic radius	B_{VR}	B site Van der Waals radius	A_{MN}	A site Mendeleev number
B_{AR}	B site atomic radius	A_{VRU}	A site UFF radius	B_{MN}	B site Mendeleev number

Table S2. Average performance (PCC) of descriptors on training sets, testing sets and independent testing sets.

Index	Training sets	Testing sets	Independent set	Calculation data
1	0.709	\	0.621	0.756
2	0.894	\	0.669	0.695
3	0.893	0.873	0.691	0.631
4	0.971	0.800	0.809	0.775
5	0.761	0.779	0.696	0.654
6	0.924	0.821	0.857	0.850

Table S3. Important hyper-parameters settings and meanings of MT-SISSO.

Hyper-parameters	Value	Description
desc_dim	2	The maximum number of dimensions of a descriptor
rung	2	The number of iterations in constructing feature space
opset	(+)(-)(*)(/)(exp)(log)(^-1)(^2)(^3)(sqrt)(cbrt)(|-|)	Set of operators
maxcomplexity	30	The maximum number of operators in a feature
maxfval_lb	1e-3	Features having the maximum absolute value <maxfval_lb will be discarded
maxfval_ub	1e5	Features having the maximum absolute value >maxfval_ub will be discarded
subs_sis	20000	The size of the subspace selected by SIS in each descriptor dimension
metric	RMSE	Error function
nsample	The number of samples in each dataset (separated by ",")	
nsf	Number of selected features	
ntask	The number of datasets	

Table S4 | Main hyper-parameters setting of Bayesian optimization.

Hyper-parameters	Value	Meaning
pbounds	a: [-5:5]; b: [-10:10]	Bounds of a and b
n_iter	500	Number of iterations
init_points	5	Number of initial points for random exploration

Table S5 | The result of Bayesian optimization using different target functions.

Target functions	Optimized a, b	The ratio of a and b	Optimized target function values
RBO	a = 1.77; b = 7.15	4.04	0.785
RA	a = 2.25; b = 8.89	3.95	0.801
PCC	a = 2.15; b = 8.81	4.10	0.743

Table S6 | The top 200 predicted active double perovskites and screened by descriptor (4) (a = 1, b = 4).

Ranking	Perovskites
1	$Cs_{0.5}Pr_{0.5}Ni_{1.0}O_3$
2	$Cs_{0.5}La_{0.5}Ni_{1.0}O_3$
3	$Rb_{0.5}Pr_{0.5}Ni_{1.0}O_3$
4	$Cs_{0.5}Pr_{0.5}Co_{0.25}Ni_{0.75}O_3$
5	$Rb_{0.5}Ce_{0.5}Ni_{1.0}O_3$
6	$Rb_{0.5}La_{0.5}Ni_{1.0}O_3$
7	$Cs_{0.5}La_{0.5}Co_{0.25}Ni_{0.75}O_3$
8	$K_{0.5}Pr_{0.5}Ni_{1.0}O_3$
9	$K_{0.5}Ce_{0.5}Ni_{1.0}O_3$
10	$K_{0.5}La_{0.5}Ni_{1.0}O_3$
11	$Rb_{0.5}Pr_{0.5}Co_{0.25}Ni_{0.75}O_3$
12	$Cs_{0.5}Pr_{0.5}Co_{0.5}Ni_{0.5}O_3$
13	$Rb_{0.5}Ce_{0.5}Co_{0.25}Ni_{0.75}O_3$
14	$Cs_{0.5}Pr_{0.5}Fe_{0.25}Ni_{0.75}O_3$
15	$Rb_{0.5}La_{0.5}Co_{0.25}Ni_{0.75}O_3$
16	$Cs_{0.5}La_{0.5}Co_{0.5}Ni_{0.5}O_3$
17	$K_{0.5}Pr_{0.5}Co_{0.25}Ni_{0.75}O_3$
18	$Cs_{0.5}La_{0.5}Fe_{0.25}Ni_{0.75}O_3$
19	$K_{0.5}Ce_{0.5}Co_{0.25}Ni_{0.75}O_3$
20	$K_{0.5}La_{0.5}Co_{0.25}Ni_{0.75}O_3$
21	$Rb_{0.5}Pr_{0.5}Co_{0.5}Ni_{0.5}O_3$
22	$Ba_{1.0}Ni_{1.0}O_3$
23	$Rb_{0.5}Pr_{0.5}Fe_{0.25}Ni_{0.75}O_3$
24	$Cs_{0.5}Pr_{0.5}Co_{0.75}Ni_{0.25}O_3$
25	$Rb_{0.5}Ce_{0.5}Co_{0.5}Ni_{0.5}O_3$
26	$Cs_{0.5}Pr_{0.5}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
27	$Rb_{0.5}La_{0.5}Co_{0.5}Ni_{0.5}O_3$
28	$Rb_{0.5}Ce_{0.5}Fe_{0.25}Ni_{0.75}O_3$
29	$Cs_{0.5}Pr_{0.5}Mn_{0.25}Ni_{0.75}O_3$
30	$Rb_{0.5}La_{0.5}Fe_{0.25}Ni_{0.75}O_3$
31	$Cs_{0.5}La_{0.5}Co_{0.75}Ni_{0.25}O_3$
32	$K_{0.5}Pr_{0.5}Co_{0.5}Ni_{0.5}O_3$
33	$Cs_{0.5}La_{0.5}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
34	$K_{0.5}Pr_{0.5}Fe_{0.25}Ni_{0.75}O_3$
35	$K_{0.5}Ce_{0.5}Co_{0.5}Ni_{0.5}O_3$
36	$Sr_{0.25}Ba_{0.75}Ni_{1.0}O_3$
37	$K_{0.5}La_{0.5}Co_{0.5}Ni_{0.5}O_3$
38	$Cs_{0.5}La_{0.5}Mn_{0.25}Ni_{0.75}O_3$
39	$K_{0.5}Ce_{0.5}Fe_{0.25}Ni_{0.75}O_3$
40	$K_{0.5}La_{0.5}Fe_{0.25}Ni_{0.75}O_3$

41	$Rb_{0.5}Pr_{0.5}Co_{0.75}Ni_{0.25}O_3$
42	$Ba_{1.0}Co_{0.25}Ni_{0.75}O_3$
43	$Rb_{0.5}Pr_{0.5}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
44	$Cs_{0.5}Pr_{0.5}Co_{1.0}O_3$
45	$Ca_{0.25}Ba_{0.75}Ni_{1.0}O_3$
46	$Rb_{0.5}Ce_{0.5}Co_{0.75}Ni_{0.25}O_3$
47	$Sr_{0.5}Ba_{0.5}Ni_{1.0}O_3$
48	$Cs_{0.5}Pr_{0.5}Fe_{0.25}Co_{0.5}Ni_{0.25}O_3$
49	$Rb_{0.5}La_{0.5}Co_{0.75}Ni_{0.25}O_3$
50	$Rb_{0.5}Pr_{0.5}Mn_{0.25}Ni_{0.75}O_3$
51	$Rb_{0.5}Ce_{0.5}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
52	$Cs_{0.5}Pr_{0.5}Mn_{0.25}Co_{0.25}Ni_{0.5}O_3$
53	$Cs_{0.5}Pr_{0.5}Fe_{0.5}Ni_{0.5}O_3$
54	$Rb_{0.5}La_{0.5}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
55	$K_{0.5}Pr_{0.5}Co_{0.75}Ni_{0.25}O_3$
56	$Cs_{0.5}La_{0.5}Co_{1.0}O_3$
57	$Rb_{0.5}Ce_{0.5}Mn_{0.25}Ni_{0.75}O_3$
58	$Cs_{0.5}La_{0.5}Fe_{0.25}Co_{0.5}Ni_{0.25}O_3$
59	$K_{0.5}Pr_{0.5}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
60	$Rb_{0.5}La_{0.5}Mn_{0.25}Ni_{0.75}O_3$
61	$K_{0.5}Ce_{0.5}Co_{0.75}Ni_{0.25}O_3$
62	$Sr_{0.25}Ba_{0.75}Co_{0.25}Ni_{0.75}O_3$
63	$Cs_{0.5}La_{0.5}Mn_{0.25}Co_{0.25}Ni_{0.5}O_3$
64	$Cs_{0.5}La_{0.5}Fe_{0.5}Ni_{0.5}O_3$
65	$K_{0.5}La_{0.5}Co_{0.75}Ni_{0.25}O_3$
66	$Sr_{0.75}Ba_{0.25}Ni_{1.0}O_3$
67	$K_{0.5}Pr_{0.5}Mn_{0.25}Ni_{0.75}O_3$
68	$K_{0.5}Ce_{0.5}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
69	$K_{0.5}La_{0.5}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
70	$K_{0.5}Ce_{0.5}Mn_{0.25}Ni_{0.75}O_3$
71	$Rb_{0.5}Pr_{0.5}Co_{1.0}O_3$
72	$Ba_{1.0}Co_{0.5}Ni_{0.5}O_3$
73	$K_{0.5}La_{0.5}Mn_{0.25}Ni_{0.75}O_3$
74	$Rb_{0.5}Pr_{0.5}Fe_{0.25}Co_{0.5}Ni_{0.25}O_3$
75	$Ca_{0.5}Ba_{0.5}Ni_{1.0}O_3$
76	$Ca_{0.25}Ba_{0.75}Co_{0.25}Ni_{0.75}O_3$
77	$Rb_{0.5}Ce_{0.5}Co_{1.0}O_3$
78	$Sr_{0.5}Ba_{0.5}Co_{0.25}Ni_{0.75}O_3$
79	$Ba_{1.0}Fe_{0.25}Ni_{0.75}O_3$
80	$Rb_{0.5}Pr_{0.5}Fe_{0.5}Ni_{0.5}O_3$
81	$Rb_{0.5}Pr_{0.5}Mn_{0.25}Co_{0.25}Ni_{0.5}O_3$
82	$Sr_{1.0}Ni_{1.0}O_3$
83	$Cs_{0.5}Pr_{0.5}Fe_{0.25}Co_{0.75}O_3$
84	$Rb_{0.5}La_{0.5}Co_{1.0}O_3$

85	$Rb_{0.5}Ce_{0.5}Fe_{0.25}Co_{0.5}Ni_{0.25}O_3$
86	$Cs_{0.5}Pr_{0.5}Mn_{0.25}Co_{0.5}Ni_{0.25}O_3$
87	$Cs_{0.5}Pr_{0.5}Fe_{0.5}Co_{0.25}Ni_{0.25}O_3$
88	$Rb_{0.5}La_{0.5}Fe_{0.25}Co_{0.5}Ni_{0.25}O_3$
89	$Rb_{0.5}Ce_{0.5}Fe_{0.5}Ni_{0.5}O_3$
90	$Rb_{0.5}Ce_{0.5}Mn_{0.25}Co_{0.25}Ni_{0.5}O_3$
91	$Cs_{0.5}Pr_{0.5}Mn_{0.25}Fe_{0.25}Ni_{0.5}O_3$
92	$K_{0.5}Pr_{0.5}Co_{1.0}O_3$
93	$Rb_{0.5}La_{0.5}Fe_{0.5}Ni_{0.5}O_3$
94	$Rb_{0.5}La_{0.5}Mn_{0.25}Co_{0.25}Ni_{0.5}O_3$
95	$K_{0.5}Pr_{0.5}Fe_{0.25}Co_{0.5}Ni_{0.25}O_3$
96	$Cs_{0.5}La_{0.5}Fe_{0.25}Co_{0.75}O_3$
97	$Sr_{0.25}Ba_{0.75}Co_{0.5}Ni_{0.5}O_3$
98	$K_{0.5}Ce_{0.5}Co_{1.0}O_3$
99	$Ca_{0.25}Sr_{0.75}Ni_{1.0}O_3$
100	$Cs_{0.5}La_{0.5}Mn_{0.25}Co_{0.5}Ni_{0.25}O_3$
101	$K_{0.5}Pr_{0.5}Fe_{0.5}Ni_{0.5}O_3$
102	$Cs_{0.5}La_{0.5}Fe_{0.5}Co_{0.25}Ni_{0.25}O_3$
103	$K_{0.5}Pr_{0.5}Mn_{0.25}Co_{0.25}Ni_{0.5}O_3$
104	$Sr_{0.75}Ba_{0.25}Co_{0.25}Ni_{0.75}O_3$
105	$K_{0.5}La_{0.5}Co_{1.0}O_3$
106	$K_{0.5}Ce_{0.5}Fe_{0.25}Co_{0.5}Ni_{0.25}O_3$
107	$Sr_{0.25}Ba_{0.75}Fe_{0.25}Ni_{0.75}O_3$
108	$Cs_{0.5}La_{0.5}Mn_{0.25}Fe_{0.25}Ni_{0.5}O_3$
109	$K_{0.5}La_{0.5}Fe_{0.25}Co_{0.5}Ni_{0.25}O_3$
110	$K_{0.5}Ce_{0.5}Fe_{0.5}Ni_{0.5}O_3$
111	$K_{0.5}Ce_{0.5}Mn_{0.25}Co_{0.25}Ni_{0.5}O_3$
112	$K_{0.5}La_{0.5}Fe_{0.5}Ni_{0.5}O_3$
113	$K_{0.5}La_{0.5}Mn_{0.25}Co_{0.25}Ni_{0.5}O_3$
114	$Ca_{0.75}Ba_{0.25}Ni_{1.0}O_3$
115	$Ba_{1.0}Co_{0.75}Ni_{0.25}O_3$
116	$Ca_{0.5}Sr_{0.5}Ni_{1.0}O_3$
117	$Ca_{0.5}Ba_{0.5}Co_{0.25}Ni_{0.75}O_3$
118	$Ca_{0.25}Ba_{0.75}Co_{0.5}Ni_{0.5}O_3$
119	$Rb_{0.5}Pr_{0.5}Fe_{0.25}Co_{0.75}O_3$
120	$Sr_{0.5}Ba_{0.5}Co_{0.5}Ni_{0.5}O_3$
121	$Ba_{1.0}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
122	$Rb_{0.5}Pr_{0.5}Mn_{0.25}Co_{0.5}Ni_{0.25}O_3$
123	$Sr_{1.0}Co_{0.25}Ni_{0.75}O_3$
124	$Rb_{0.5}Pr_{0.5}Fe_{0.5}Co_{0.25}Ni_{0.25}O_3$
125	$Ca_{0.25}Ba_{0.75}Fe_{0.25}Ni_{0.75}O_3$
126	$Rb_{0.5}Pr_{0.5}Mn_{0.25}Fe_{0.25}Ni_{0.5}O_3$
127	$Rb_{0.5}Ce_{0.5}Fe_{0.25}Co_{0.75}O_3$
128	$Sr_{0.5}Ba_{0.5}Fe_{0.25}Ni_{0.75}O_3$

129	$Cs_{0.5}Pr_{0.5}Mn_{0.25}Co_{0.75}O_3$
130	$Ba_{1.0}Mn_{0.25}Ni_{0.75}O_3$
131	$Cs_{0.5}Pr_{0.5}Fe_{0.5}Co_{0.5}O_3$
132	$Rb_{0.5}Ce_{0.5}Mn_{0.25}Co_{0.5}Ni_{0.25}O_3$
133	$Cs_{0.5}Pr_{0.5}Mn_{0.5}Ni_{0.5}O_3$
134	$Cs_{0.5}Pr_{0.5}Mn_{0.25}Fe_{0.25}Co_{0.25}Ni_{0.25}O_3$
135	$Rb_{0.5}La_{0.5}Fe_{0.25}Co_{0.75}O_3$
136	$Rb_{0.5}Ce_{0.5}Fe_{0.5}Co_{0.25}Ni_{0.25}O_3$
137	$Ca_{0.75}Sr_{0.25}Ni_{1.0}O_3$
138	$Cs_{0.5}Pr_{0.5}Fe_{0.75}Ni_{0.25}O_3$
139	$Rb_{0.5}Ce_{0.5}Mn_{0.25}Fe_{0.25}Ni_{0.5}O_3$
140	$Rb_{0.5}La_{0.5}Mn_{0.25}Co_{0.5}Ni_{0.25}O_3$
141	$Rb_{0.5}La_{0.5}Fe_{0.5}Co_{0.25}Ni_{0.25}O_3$
142	$K_{0.5}Pr_{0.5}Fe_{0.25}Co_{0.75}O_3$
143	$Sr_{0.25}Ba_{0.75}Co_{0.75}Ni_{0.25}O_3$
144	$Rb_{0.5}La_{0.5}Mn_{0.25}Fe_{0.25}Ni_{0.5}O_3$
145	$Ca_{0.25}Sr_{0.75}Co_{0.25}Ni_{0.75}O_3$
146	$Cs_{0.5}La_{0.5}Mn_{0.25}Co_{0.75}O_3$
147	$K_{0.5}Pr_{0.5}Mn_{0.25}Co_{0.5}Ni_{0.25}O_3$
148	$Sr_{0.75}Ba_{0.25}Co_{0.5}Ni_{0.5}O_3$
149	$K_{0.5}Pr_{0.5}Fe_{0.5}Co_{0.25}Ni_{0.25}O_3$
150	$Cs_{0.5}La_{0.5}Fe_{0.5}Co_{0.5}O_3$
151	$Cs_{0.5}La_{0.5}Mn_{0.25}Fe_{0.25}Co_{0.25}Ni_{0.25}O_3$
152	$Cs_{0.5}La_{0.5}Mn_{0.5}Ni_{0.5}O_3$
153	$Sr_{0.25}Ba_{0.75}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
154	$K_{0.5}Ce_{0.5}Fe_{0.25}Co_{0.75}O_3$
155	$K_{0.5}Pr_{0.5}Mn_{0.25}Fe_{0.25}Ni_{0.5}O_3$
156	$Cs_{0.5}La_{0.5}Fe_{0.75}Ni_{0.25}O_3$
157	$Ca_{1.0}Ni_{1.0}O_3$
158	$Sr_{0.75}Ba_{0.25}Fe_{0.25}Ni_{0.75}O_3$
159	$K_{0.5}Ce_{0.5}Mn_{0.25}Co_{0.5}Ni_{0.25}O_3$
160	$K_{0.5}La_{0.5}Fe_{0.25}Co_{0.75}O_3$
161	$K_{0.5}Ce_{0.5}Fe_{0.5}Co_{0.25}Ni_{0.25}O_3$
162	$Sr_{0.25}Ba_{0.75}Mn_{0.25}Ni_{0.75}O_3$
163	$Ca_{0.75}Ba_{0.25}Co_{0.25}Ni_{0.75}O_3$
164	$K_{0.5}La_{0.5}Mn_{0.25}Co_{0.5}Ni_{0.25}O_3$
165	$K_{0.5}Ce_{0.5}Mn_{0.25}Fe_{0.25}Ni_{0.5}O_3$
166	$K_{0.5}La_{0.5}Fe_{0.5}Co_{0.25}Ni_{0.25}O_3$
167	$Ca_{0.5}Sr_{0.5}Co_{0.25}Ni_{0.75}O_3$
168	$Ba_{1.0}Co_{1.0}O_3$
169	$Ca_{0.5}Ba_{0.5}Co_{0.5}Ni_{0.5}O_3$
170	$K_{0.5}La_{0.5}Mn_{0.25}Fe_{0.25}Ni_{0.5}O_3$
171	$Ca_{0.25}Ba_{0.75}Co_{0.75}Ni_{0.25}O_3$
172	$Sr_{0.5}Ba_{0.5}Co_{0.75}Ni_{0.25}O_3$

173	$Ba_{1.0}Fe_{0.25}Co_{0.5}Ni_{0.25}O_3$
174	$Rb_{0.5}Pr_{0.5}Mn_{0.25}Co_{0.75}O_3$
175	$Sr_{1.0}Co_{0.5}Ni_{0.5}O_3$
176	$Ca_{0.5}Ba_{0.5}Fe_{0.25}Ni_{0.75}O_3$
177	$Ca_{0.25}Ba_{0.75}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
178	$Rb_{0.5}Pr_{0.5}Fe_{0.5}Co_{0.5}O_3$
179	$Rb_{0.5}Pr_{0.5}Mn_{0.25}Fe_{0.25}Co_{0.25}Ni_{0.25}O_3$
180	$Ba_{1.0}Fe_{0.5}Ni_{0.5}O_3$
181	$Rb_{0.5}Pr_{0.5}Mn_{0.5}Ni_{0.5}O_3$
182	$Sr_{0.5}Ba_{0.5}Fe_{0.25}Co_{0.25}Ni_{0.5}O_3$
183	$Ba_{1.0}Mn_{0.25}Co_{0.25}Ni_{0.5}O_3$
184	$Sr_{1.0}Fe_{0.25}Ni_{0.75}O_3$
185	$Rb_{0.5}Pr_{0.5}Fe_{0.75}Ni_{0.25}O_3$
186	$Cs_{0.5}Pr_{0.5}Mn_{0.5}Co_{0.25}Ni_{0.25}O_3$
187	$Rb_{0.5}Ce_{0.5}Mn_{0.25}Co_{0.75}O_3$
188	$Cs_{0.5}Pr_{0.5}Mn_{0.25}Fe_{0.25}Co_{0.5}O_3$
189	$Ca_{0.75}Sr_{0.25}Co_{0.25}Ni_{0.75}O_3$
190	$Rb_{0.5}Ce_{0.5}Fe_{0.5}Co_{0.5}O_3$
191	$Ca_{0.25}Ba_{0.75}Mn_{0.25}Ni_{0.75}O_3$
192	$Rb_{0.5}Ce_{0.5}Mn_{0.25}Fe_{0.25}Co_{0.25}Ni_{0.25}O_3$
193	$Cs_{0.5}Pr_{0.5}Mn_{0.25}Fe_{0.5}Ni_{0.25}O_3$
194	$Rb_{0.5}La_{0.5}Mn_{0.25}Co_{0.75}O_3$
195	$Rb_{0.5}Ce_{0.5}Mn_{0.5}Ni_{0.5}O_3$
196	$Cs_{0.5}Pr_{0.5}Fe_{0.75}Co_{0.25}O_3$
197	$Sr_{0.5}Ba_{0.5}Mn_{0.25}Ni_{0.75}O_3$
198	$Rb_{0.5}Ce_{0.5}Fe_{0.75}Ni_{0.25}O_3$
199	$Rb_{0.5}La_{0.5}Fe_{0.5}Co_{0.5}O_3$
200	$Rb_{0.5}La_{0.5}Mn_{0.25}Fe_{0.25}Co_{0.25}Ni_{0.25}O_3$

Table S7 | The top 100 predicted active double perovskites with Ir and Ni in B-site screened by descriptor (4) (a = 1, b = 4)

Ranking	Perovskites
1	$Pr_{0.5}Cs_{0.5}Ir_{0.3}Ni_{0.7}O_3$
2	$Pr_{0.5}Cs_{0.5}Ir_{0.4}Ni_{0.6}O_3$
3	$Pr_{0.5}Cs_{0.5}Ir_{0.2}Ni_{0.8}O_3$
4	$Pr_{0.5}Cs_{0.5}Ir_{0.5}Ni_{0.5}O_3$
5	$Pr_{0.5}Cs_{0.5}Ir_{0.1}Ni_{0.9}O_3$
6	$Pr_{0.5}Cs_{0.5}Ir_{0.6}Ni_{0.4}O_3$
7	$Nd_{0.5}Cs_{0.5}Ir_{0.3}Ni_{0.7}O_3$
8	$Nd_{0.5}Cs_{0.5}Ir_{0.4}Ni_{0.6}O_3$
9	$Pr_{0.5}Cs_{0.5}Ni1.0O_3$
10	$Nd_{0.5}Cs_{0.5}Ir_{0.2}Ni_{0.8}O_3$
11	$Ce_{0.5}Cs_{0.5}Ir_{0.3}Ni_{0.7}O_3$
12	$Ce_{0.5}Cs_{0.5}Ir_{0.4}Ni_{0.6}O_3$
13	$Nd_{0.5}Cs_{0.5}Ir_{0.5}Ni_{0.5}O_3$
14	$Ce_{0.5}Cs_{0.5}Ir_{0.2}Ni_{0.8}O_3$
15	$Nd_{0.5}Cs_{0.5}Ir_{0.1}Ni_{0.9}O_3$
16	$Pr_{0.5}Cs_{0.5}Ir_{0.7}Ni_{0.3}O_3$
17	$Ce_{0.5}Cs_{0.5}Ir_{0.5}Ni_{0.5}O_3$
18	$Ce_{0.5}Cs_{0.5}Ir_{0.1}Ni_{0.9}O_3$
19	$Nd_{0.5}Cs_{0.5}Ir_{0.6}Ni_{0.4}O_3$
20	$La_{0.5}Cs_{0.5}Ir_{0.3}Ni_{0.7}O_3$
21	$La_{0.5}Cs_{0.5}Ir_{0.4}Ni_{0.6}O_3$
22	$Nd_{0.5}Cs_{0.5}Ni1.0O_3$
23	$La_{0.5}Cs_{0.5}Ir_{0.2}Ni_{0.8}O_3$
24	$Ce_{0.5}Cs_{0.5}Ir_{0.6}Ni_{0.4}O_3$
25	$Ce_{0.5}Cs_{0.5}Ni1.0O_3$
26	$La_{0.5}Cs_{0.5}Ir_{0.5}Ni_{0.5}O_3$
27	$Pr_{0.5}Cs_{0.5}Ir_{0.8}Ni_{0.2}O_3$
28	$La_{0.5}Cs_{0.5}Ir_{0.1}Ni_{0.9}O_3$
29	$Nd_{0.5}Cs_{0.5}Ir_{0.7}Ni_{0.3}O_3$
30	$Ce_{0.5}Cs_{0.5}Ir_{0.7}Ni_{0.3}O_3$
31	$La_{0.5}Cs_{0.5}Ir_{0.6}Ni_{0.4}O_3$
32	$La_{0.5}Cs_{0.5}Ni1.0O_3$
33	$Nd_{0.5}Cs_{0.5}Ir_{0.8}Ni_{0.2}O_3$
34	$Pr_{0.5}Cs_{0.5}Ir_{0.9}Ni_{0.1}O_3$
35	$La_{0.5}Cs_{0.5}Ir_{0.7}Ni_{0.3}O_3$
36	$Ce_{0.5}Cs_{0.5}Ir_{0.8}Ni_{0.2}O_3$
37	$La_{0.5}Cs_{0.5}Ir_{0.8}Ni_{0.2}O_3$
38	$Nd_{0.5}Cs_{0.5}Ir_{0.9}Ni_{0.1}O_3$
39	$Pr_{0.5}Cs_{0.5}Ir1.0O_3$
40	$Ce_{0.5}Cs_{0.5}Ir_{0.9}Ni_{0.1}O_3$
41	$Rb_{0.5}Pr_{0.5}Ir_{0.3}Ni_{0.7}O_3$

42	$Rb_{0.5}Pr_{0.5}Ir_{0.4}Ni_{0.6}O_3$
43	$La_{0.5}Cs_{0.5}Ir_{0.9}Ni_{0.1}O_3$
44	$Rb_{0.5}Pr_{0.5}Ir_{0.2}Ni_{0.8}O_3$
45	$Nd_{0.5}Cs_{0.5}Ir1.0O_3$
46	$Rb_{0.5}Pr_{0.5}Ir_{0.5}Ni_{0.5}O_3$
47	$Rb_{0.5}Pr_{0.5}Ir_{0.1}Ni_{0.9}O_3$
48	$Ce_{0.5}Cs_{0.5}Ir1.0O_3$
49	$Rb_{0.5}Pr_{0.5}Ir_{0.6}Ni_{0.4}O_3$
50	$Nd_{0.5}Rb_{0.5}Ir_{0.3}Ni_{0.7}O_3$
51	$Nd_{0.5}Rb_{0.5}Ir_{0.4}Ni_{0.6}O_3$
52	$Rb_{0.5}Pr_{0.5}Ni1.0O_3$
53	$Nd_{0.5}Rb_{0.5}Ir_{0.2}Ni_{0.8}O_3$
54	$Rb_{0.5}Ce_{0.5}Ir_{0.3}Ni_{0.7}O_3$
55	$Rb_{0.5}Ce_{0.5}Ir_{0.4}Ni_{0.6}O_3$
56	$La_{0.5}Cs_{0.5}Ir1.0O_3$
57	$Rb_{0.5}Ce_{0.5}Ir_{0.2}Ni_{0.8}O_3$
58	$Nd_{0.5}Rb_{0.5}Ir_{0.5}Ni_{0.5}O_3$
59	$Nd_{0.5}Rb_{0.5}Ir_{0.1}Ni_{0.9}O_3$
60	$Rb_{0.5}Pr_{0.5}Ir_{0.7}Ni_{0.3}O_3$
61	$Rb_{0.5}Ce_{0.5}Ir_{0.5}Ni_{0.5}O_3$
62	$Rb_{0.5}Ce_{0.5}Ir_{0.1}Ni_{0.9}O_3$
63	$Cs_{0.5}Pr_{0.5}Co_{0.2}5Ni_{0.7}5O_3$
64	$Nd_{0.5}Rb_{0.5}Ir_{0.6}Ni_{0.4}O_3$
65	$Rb_{0.5}La_{0.5}Ir_{0.3}Ni_{0.7}O_3$
66	$Rb_{0.5}La_{0.5}Ir_{0.4}Ni_{0.6}O_3$
67	$Rb_{0.5}La_{0.5}Ir_{0.2}Ni_{0.8}O_3$
68	$Nd_{0.5}Rb_{0.5}Ni1.0O_3$
69	$Rb_{0.5}Ce_{0.5}Ir_{0.6}Ni_{0.4}O_3$
70	$Rb_{0.5}Ce_{0.5}Ni1.0O_3$
71	$Rb_{0.5}La_{0.5}Ir_{0.5}Ni_{0.5}O_3$
72	$Rb_{0.5}La_{0.5}Ir_{0.1}Ni_{0.9}O_3$
73	$Rb_{0.5}Pr_{0.5}Ir_{0.8}Ni_{0.2}O_3$
74	$Nd_{0.5}Rb_{0.5}Ir_{0.7}Ni_{0.3}O_3$
75	$Rb_{0.5}Ce_{0.5}Ir_{0.7}Ni_{0.3}O_3$
76	$Rb_{0.5}La_{0.5}Ir_{0.6}Ni_{0.4}O_3$
77	$Rb_{0.5}La_{0.5}Ni1.0O_3$
78	$K_{0.5}Pr_{0.5}Ir_{0.3}Ni_{0.7}O_3$
79	$K_{0.5}Pr_{0.5}Ir_{0.4}Ni_{0.6}O_3$
80	$K_{0.5}Pr_{0.5}Ir_{0.2}Ni_{0.8}O_3$
81	$K_{0.5}Pr_{0.5}Ir_{0.5}Ni_{0.5}O_3$
82	$Nd_{0.5}Rb_{0.5}Ir_{0.8}Ni_{0.2}O_3$
83	$K_{0.5}Pr_{0.5}Ir_{0.1}Ni_{0.9}O_3$
84	$Rb_{0.5}La_{0.5}Ir_{0.7}Ni_{0.3}O_3$
85	$Rb_{0.5}Pr_{0.5}Ir_{0.9}Ni_{0.1}O_3$

86	$Rb_{0.5}Ce_{0.5}Ir_{0.8}Ni_{0.2}O_3$
87	$Cs_{0.5}La_{0.5}Co_{0.2}5Ni_{0.7}5O_3$
88	$K_{0.5}Pr_{0.5}Ir_{0.6}Ni_{0.4}O_3$
89	$K_{0.5}Nd_{0.5}Ir_{0.3}Ni_{0.7}O_3$
90	$K_{0.5}Nd_{0.5}Ir_{0.4}Ni_{0.6}O_3$
91	$K_{0.5}Nd_{0.5}Ir_{0.2}Ni_{0.8}O_3$
92	$K_{0.5}Pr_{0.5}Ni1.0O_3$
93	$K_{0.5}Ce_{0.5}Ir_{0.3}Ni_{0.7}O_3$
94	$K_{0.5}Ce_{0.5}Ir_{0.4}Ni_{0.6}O_3$
95	$K_{0.5}Ce_{0.5}Ir_{0.2}Ni_{0.8}O_3$
96	$K_{0.5}Nd_{0.5}Ir_{0.5}Ni_{0.5}O_3$
97	$Rb_{0.5}La_{0.5}Ir_{0.8}Ni_{0.2}O_3$
98	$K_{0.5}Nd_{0.5}Ir_{0.1}Ni_{0.9}O_3$
99	$Nd_{0.5}Rb_{0.5}Ir_{0.9}Ni_{0.1}O_3$
100	$K_{0.5}Pr_{0.5}Ir_{0.7}Ni_{0.3}O_3$

Table S8 | Calculated overpotentials and the Gibbs free energy differences of each step (ΔG_{O*}, ΔG_{OH*} and ΔG_{OOH*}) of 15 double perovskites.

Index	Materials	ΔG_{O*}(eV)	ΔG_{OH*}(eV)	ΔG_{OOH*}(eV)	η_{OER}(V)
1	K0.5Ce0.5Ni1.0O3	3.679	1.773	4.517	0.640
2	Cs0.5La0.5Ni1.0O3	3.091	1.220	4.205	0.675
3	$BaNiO_3$	3.409	1.426	4.580	0.752
4	$Ba_{0.75}Sr_{0.25}Ni_{1.0}O_3$	2.726	1.945	4.596	0.715
5	$BaCo_{0.25}Ni_{0.75}O_3$	4.174	2.068	4.792	0.877
6	$Ba_{0.75}Ca_{0.25}Ni_{1.0}O_3$	3.297	0.927	4.600	1.140
7	$Ba_{0.5}Sr_{0.5}Ni_{1.0}O_3$	4.090	2.011	4.674	0.848
8	$Ba_{0.25}Sr_{0.75}Ni_{1.0}O_3$	4.280	1.960	4.550	1.090
9	$BaCo_{0.5}Ni_{0.5}O_3$	3.662	1.194	3.930	1.238
10	$Ba_{0.5}Ca_{0.5}Ni_{1.0}O_3$	2.805	2.001	4.709	0.771
11	$BaFe_{0.25}Ni_{0.75}O_3$	4.161	1.792	4.652	1.139
12	$SrNiO_3$	4.366	2.025	4.591	1.112
13	$BaMn_{0.25}Ni_{0.75}O_3$	4.178	1.778	4.612	1.170
14	$SrCo_{0.5}Ni_{0.5}O_3$	3.559	1.217	4.650	1.111
15	$BaFe_{0.5}Ni_{0.5}O_3$	4.222	1.781	4.865	1.211

Supplementary Figures

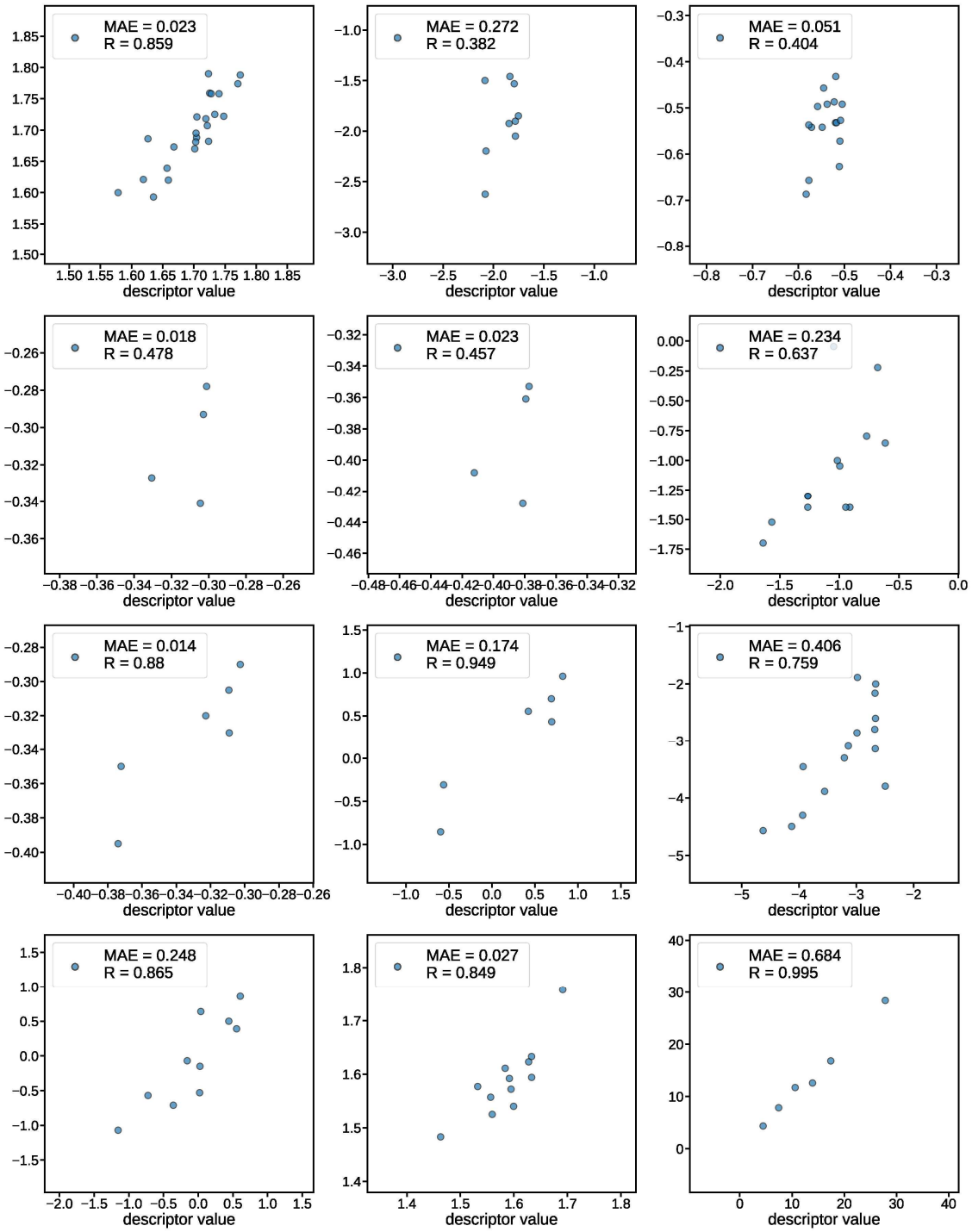

Figure S1 | Overall performance of the original best 1D descriptor on 12 datasets.

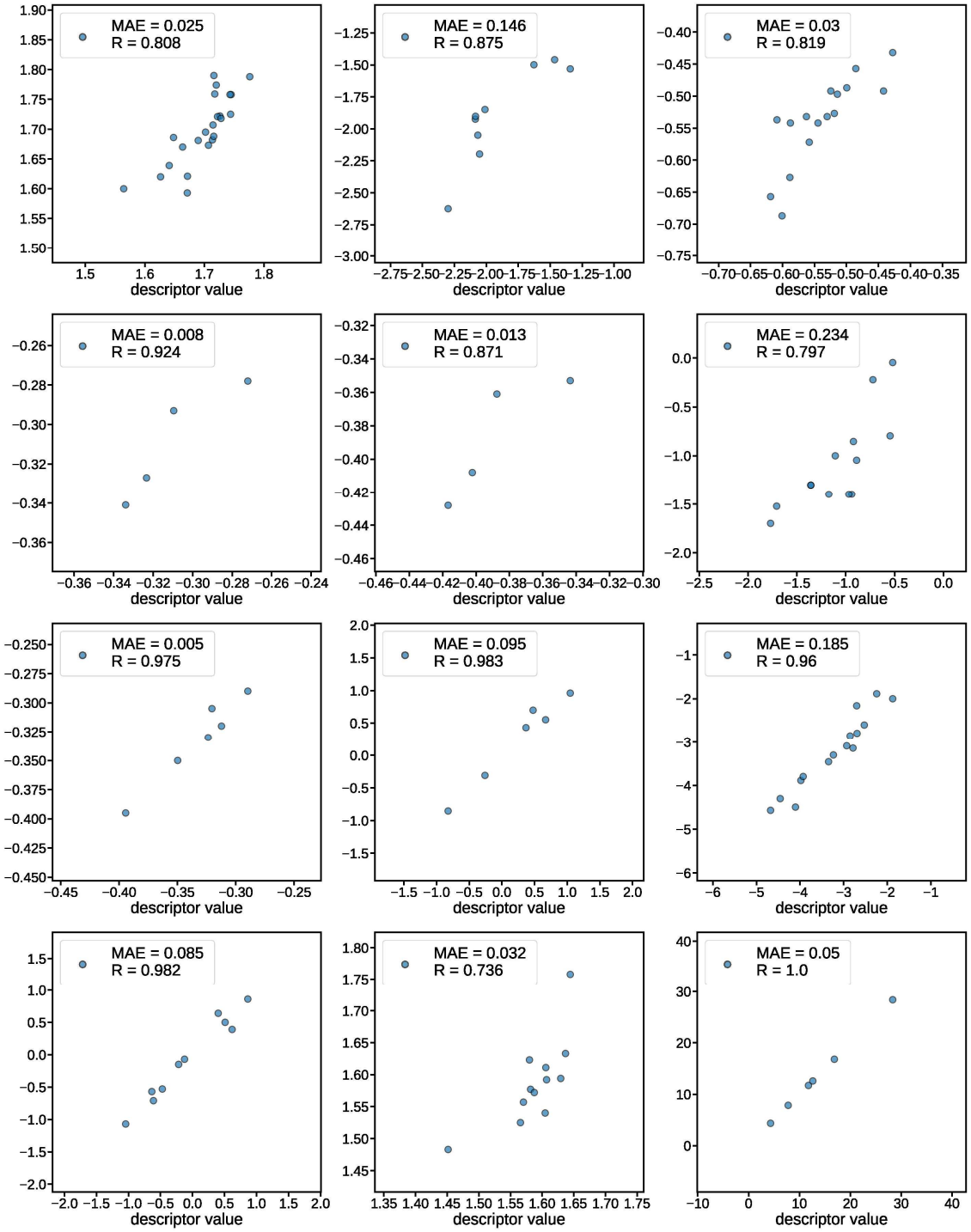

Figure S2 | Overall performance of the original best 2D descriptor on 12 datasets.

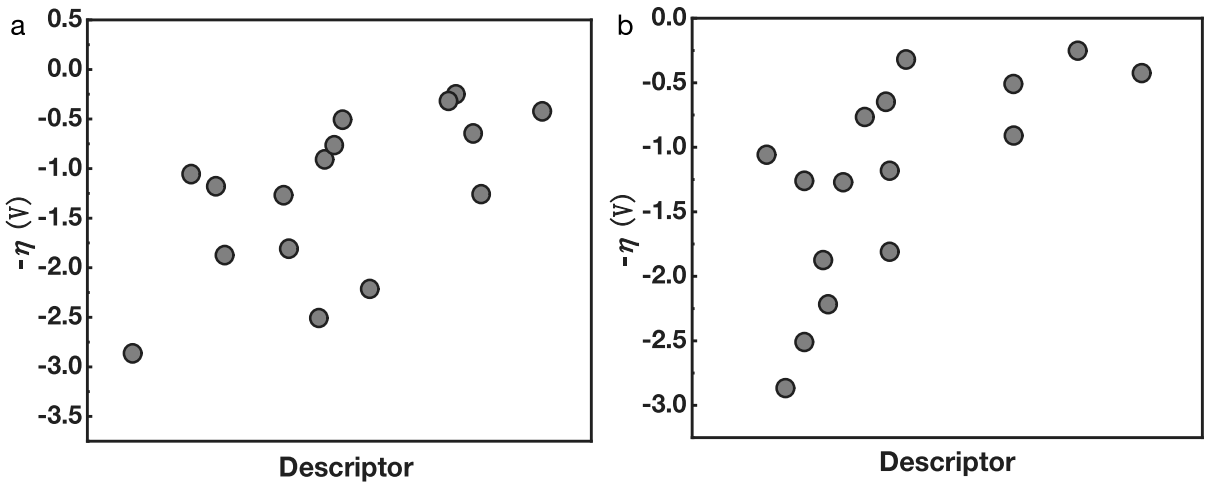

Figure S3 | The original performance of descriptors derived from 12 datasets on independent testing set. a, 1D descriptor (PCC = 0.606). b, 2D descriptor (PCC = 0.669)

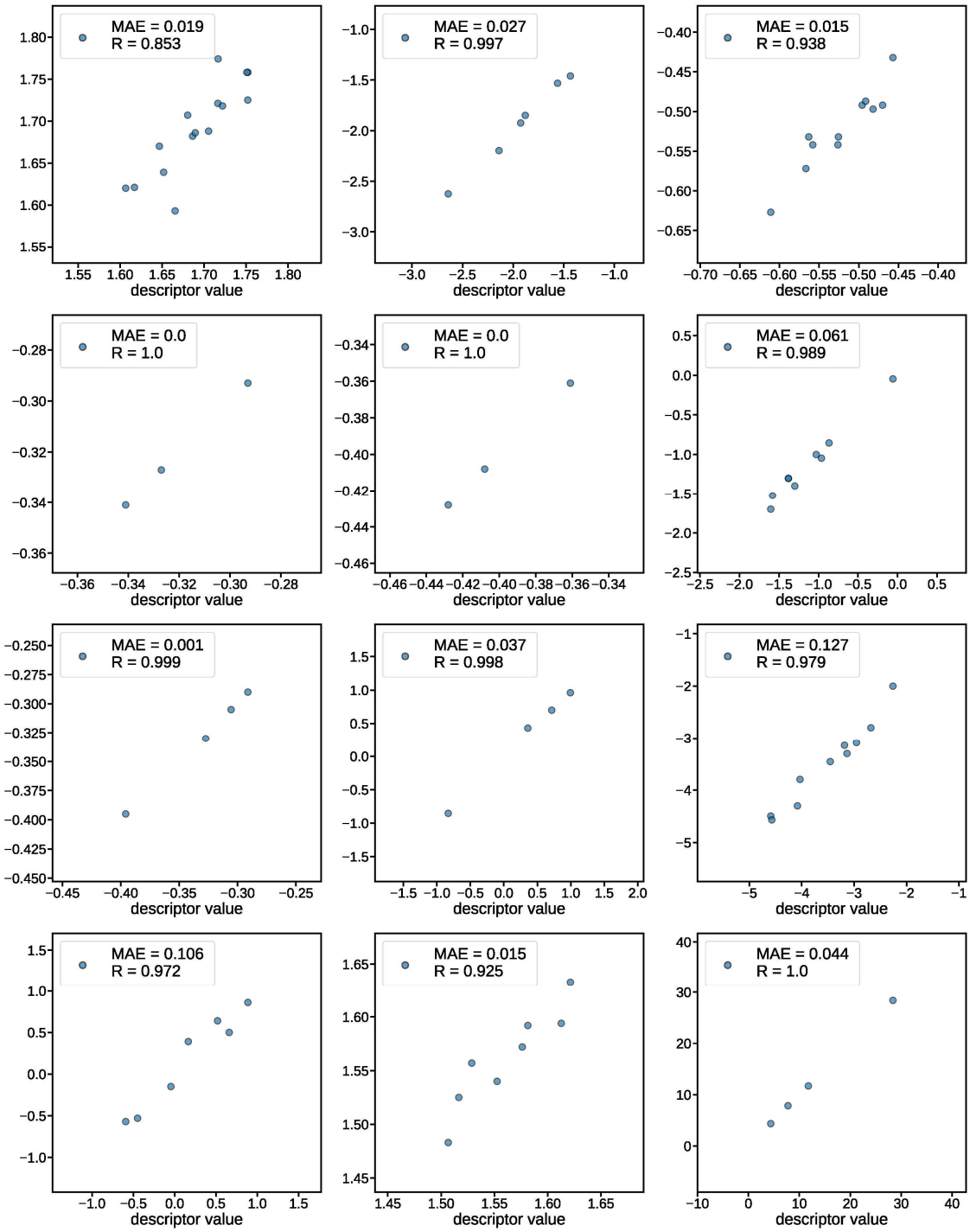

Figure S4 | Overall performance of the best 2D descriptor derived from 12 datasets after feature selection and k-fold CV

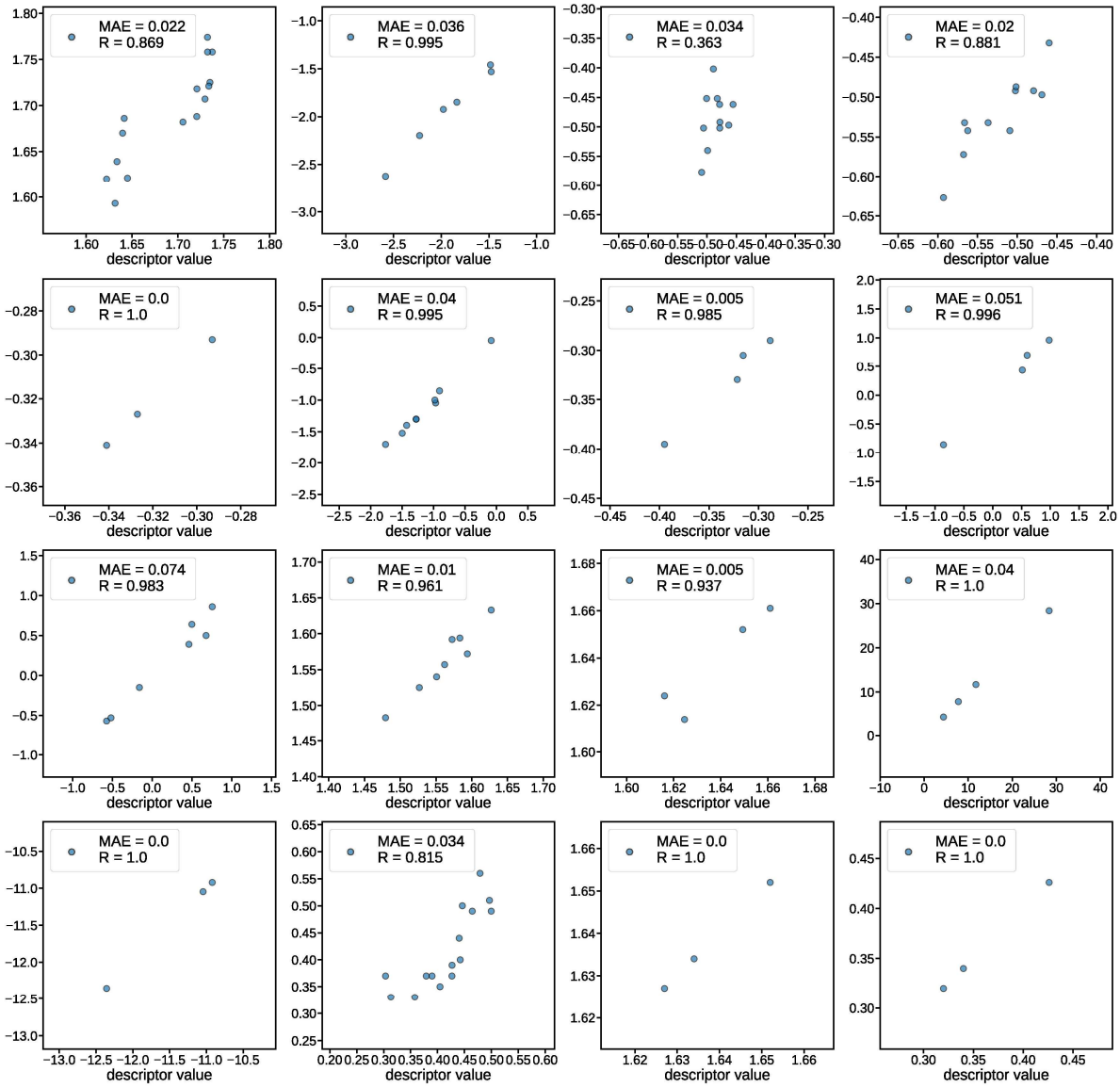

Figure S5 | Overall performance of the best 2d descriptor derived from 16 datasets after feature selection and k-fold CV.

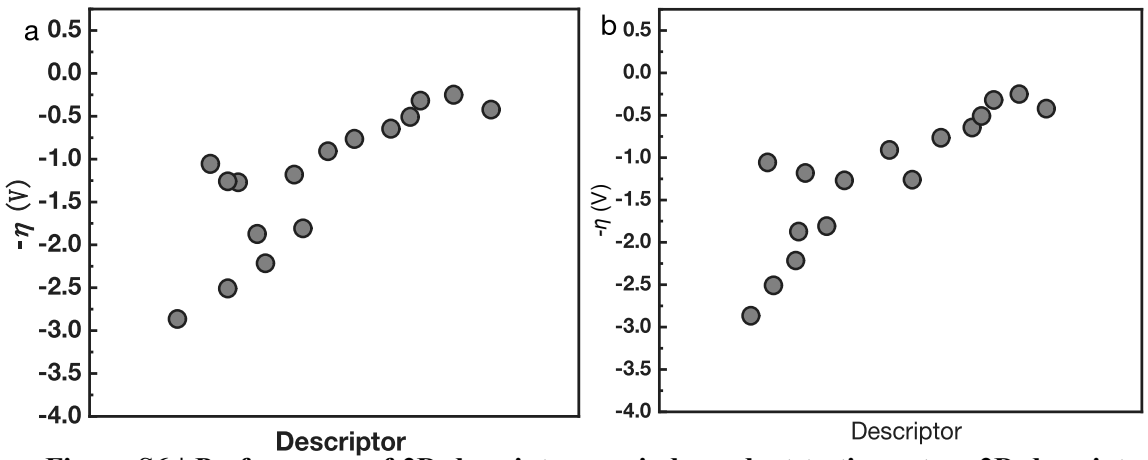

Figure S6 | Performance of 2D descriptors on independent testing set. a, 2D descriptor derived from 12 datasets (PCC = 0.80). a, 2D descriptor derived from 16 datasets (PCC = 0.86)

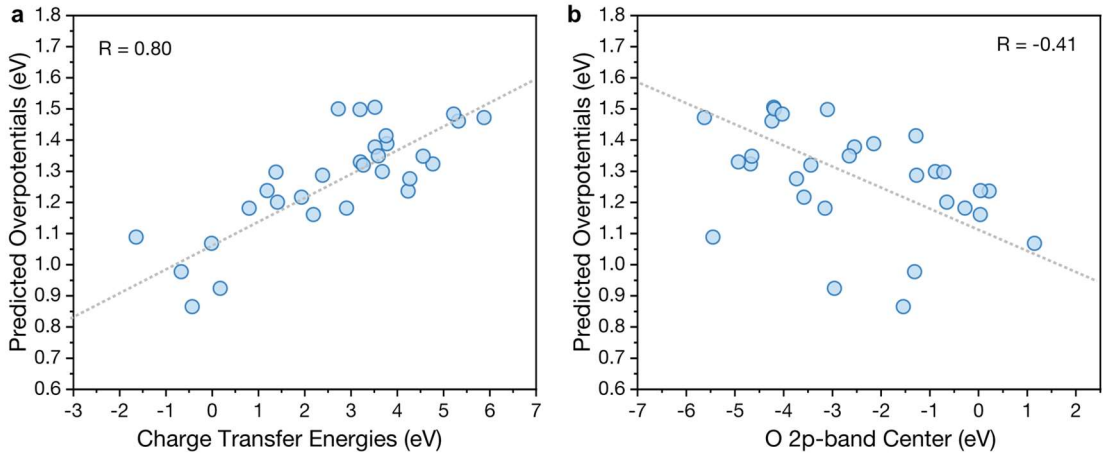

Figure S7 | Linear correlations between predicted overpotentials via descriptor 6 and (a) charge transfer energy (PCC = 0.80) and (b), the O 2*p* band center relative to the Fermi level (PCC = -0.41).

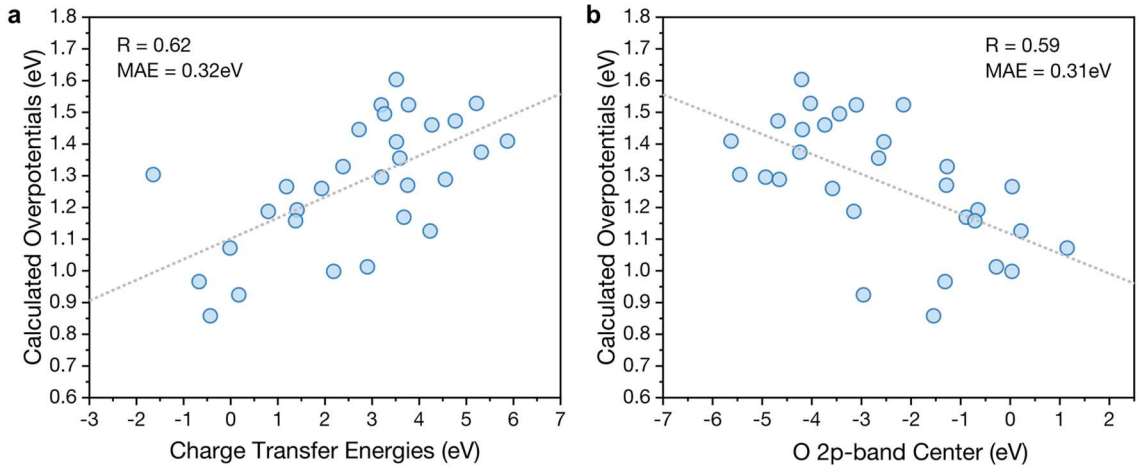

Figure S8 | Linear correlations between calculated overpotentials and (a) charge transfer energy (PCC = 0.62) and (b) the O 2*p* band center relative to the Fermi level (PCC = -0.59).

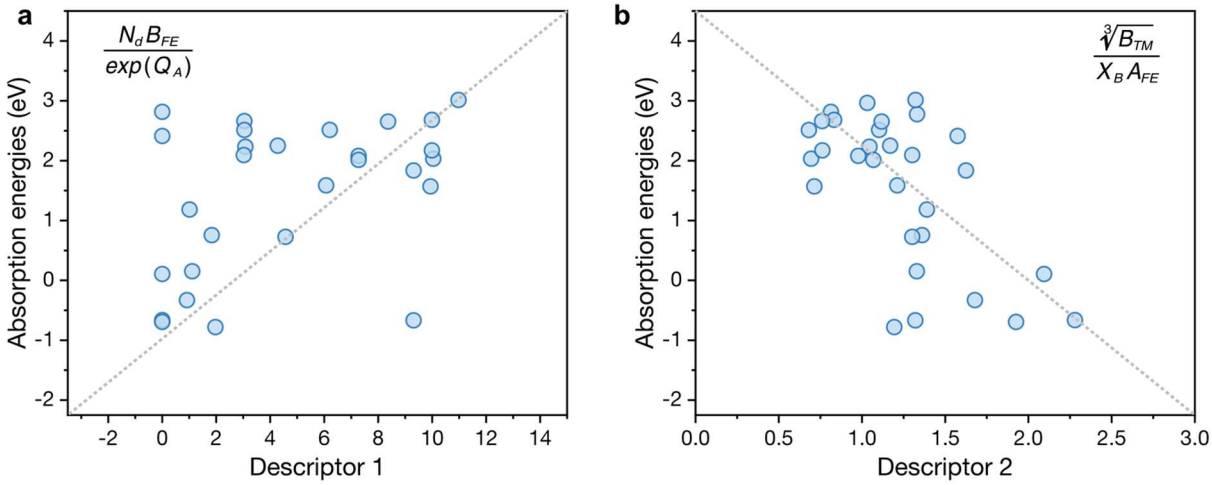

Figure S9 | Linear correlations between absorption energy of intermediates and a, the left part of descriptor 6 and b, the right part of MT-descriptor.

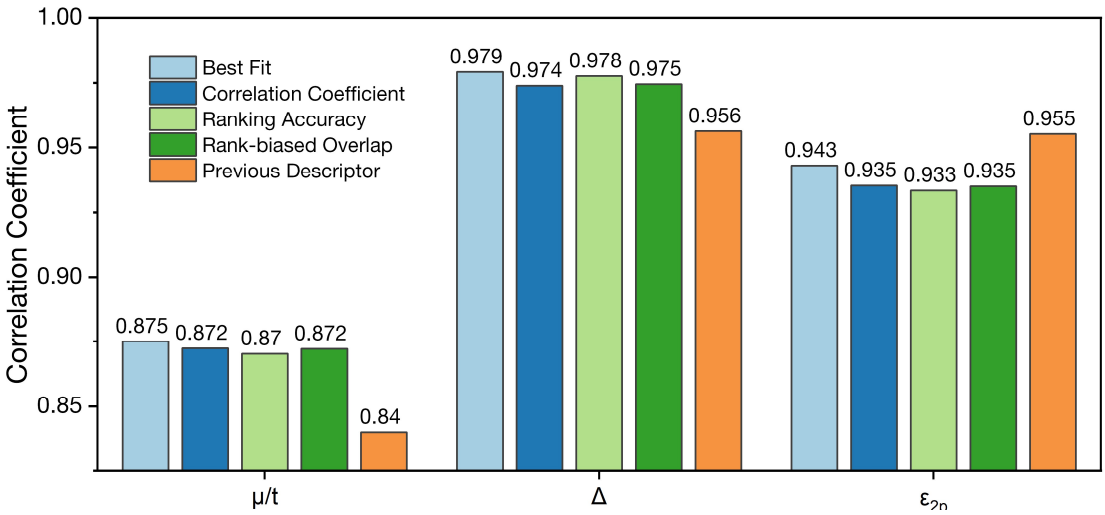

Figure S10 | Comparison between the PCC of descriptors proposed in literature and PCC of descriptors 6 using recommended parameters optimized by different target functions and best-fitting parameters on literature data.

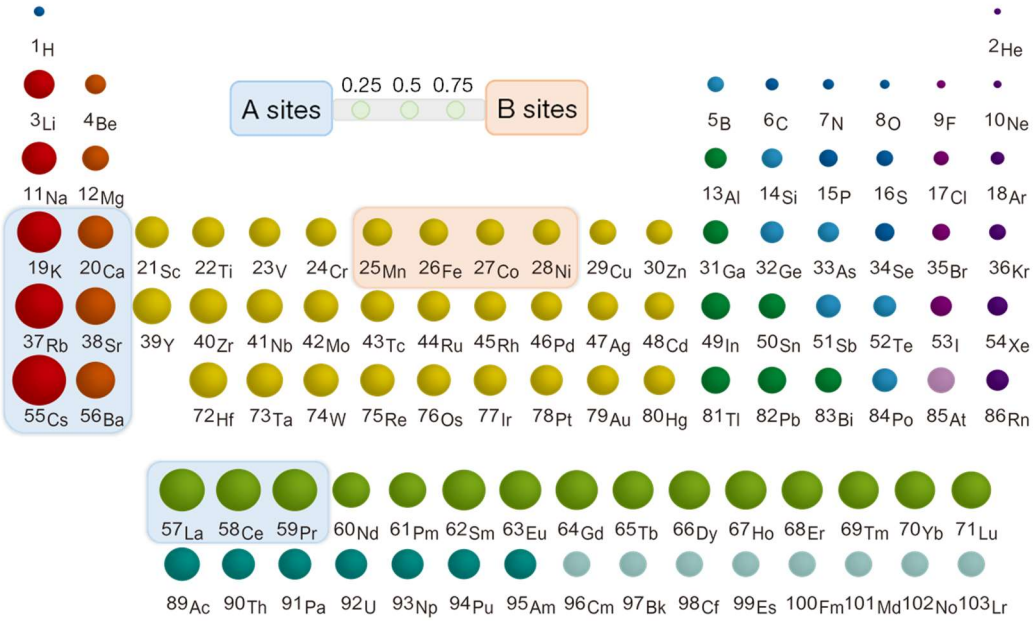

Figure S11 | Choice of A site and B site elements to screen double perovskites in the periodic table.

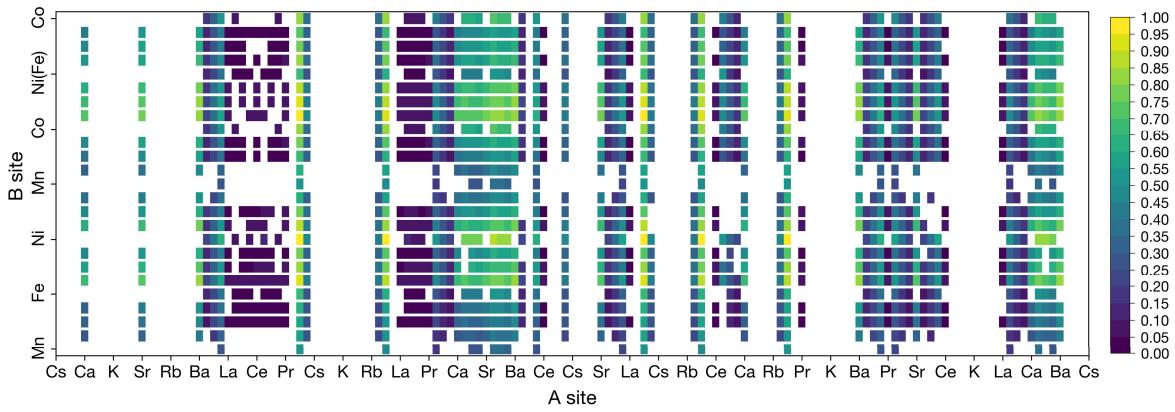

Figure S12 | The predicted activities of 3545 double perovskites using MT-descriptor with optimized parameters. The A(B) site element of a double perovskite is represented by the A(B) site label. And if the column is located at 0.25, 0.5, 0.75 position between two element labels X, Y, the A(B)-site element is $X_{0.25}Y_{0.75}$, $X_{0.5}Y_{0.5}$, $X_{0.75}Y_{0.25}$, respectively.

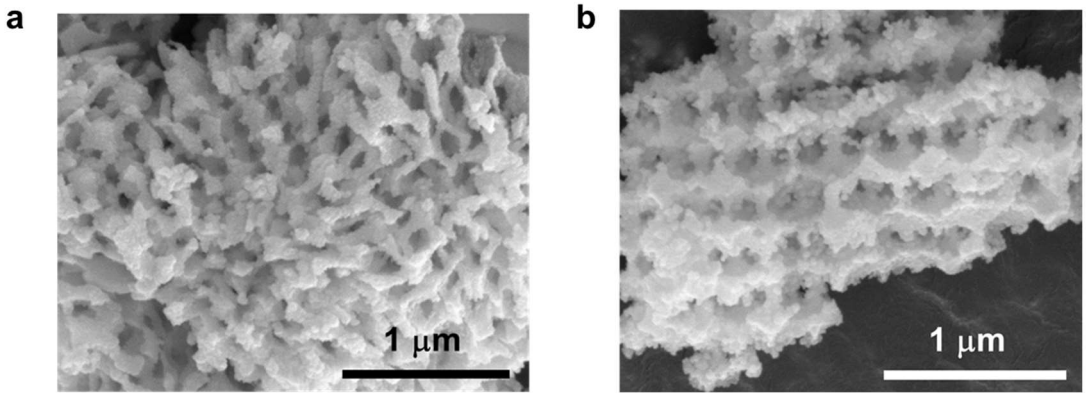

Figure S13 | SEM images of **(c)**$Cs_{0.4}La_{0.6}NiO_3$ and **(d)**$K_{0.5}Ce_{0.5}NiO_3$.

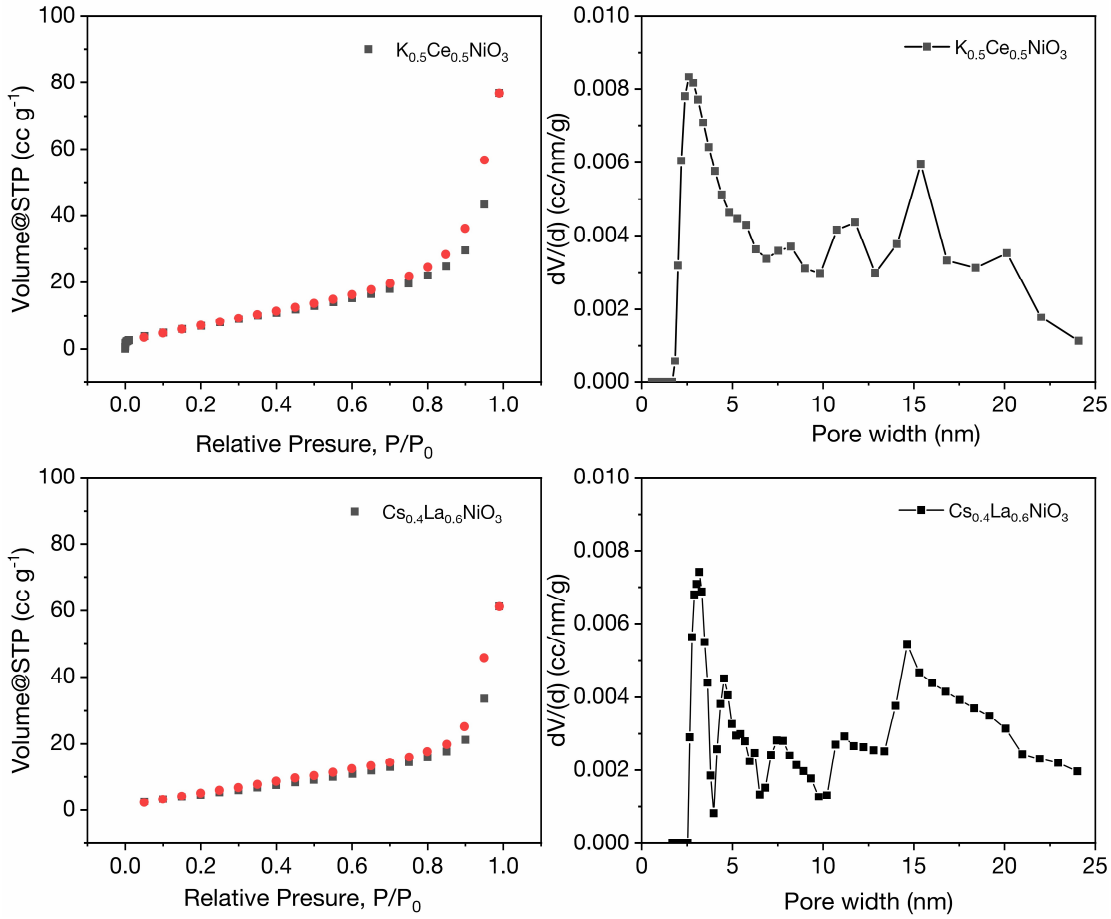

Figure S14 | (a) N_2 adsorption-desorption isotherms, (b) pore size distributions of $K_{0.5}Ce_{0.5}NiO_3$; (c) N_2 adsorption-desorption isotherms, (d) pore size distributions of $Cs_{0.5}La_{0.5}NiO_3$.

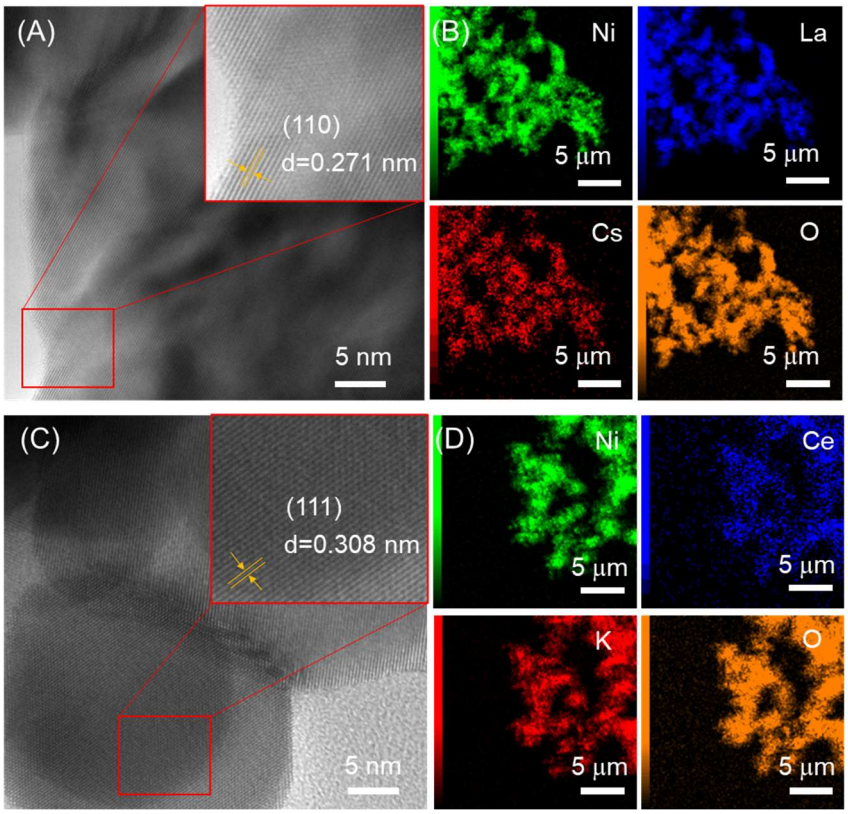

Figure S15 | (a) HR-TEM image and (b) EDX mappings of $Cs_{0.4}La_{0.6}NiO_3$ sample. (c) HR-TEM image and (d) EDX mappings of $K_{0.5}Ce_{0.5}NiO_3$ sample.

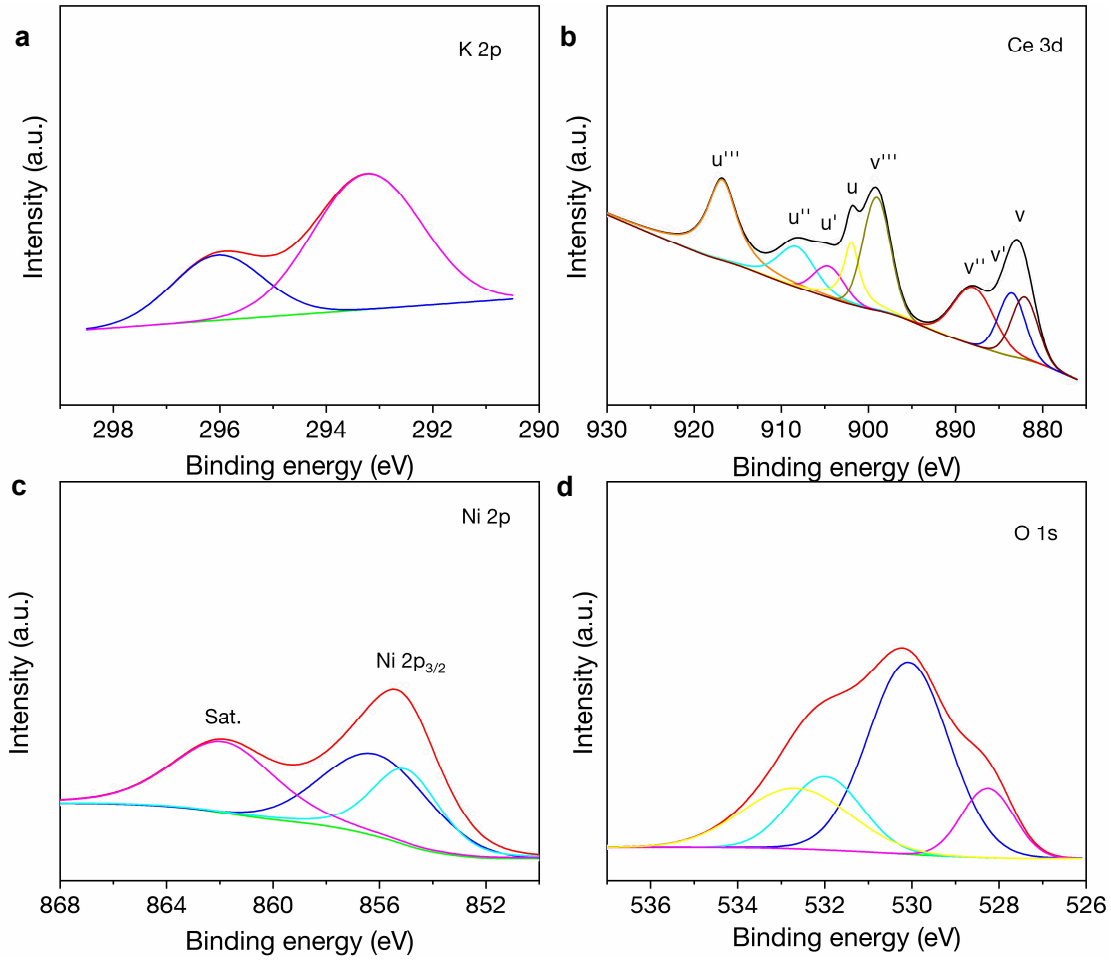

Figure S16 | XPS spectra for (a) K 2p, (b) Ce 3d, (c) Ni 2p, (d) O 1s of $K_{0.5}Ce_{0.5}NiO_3$.

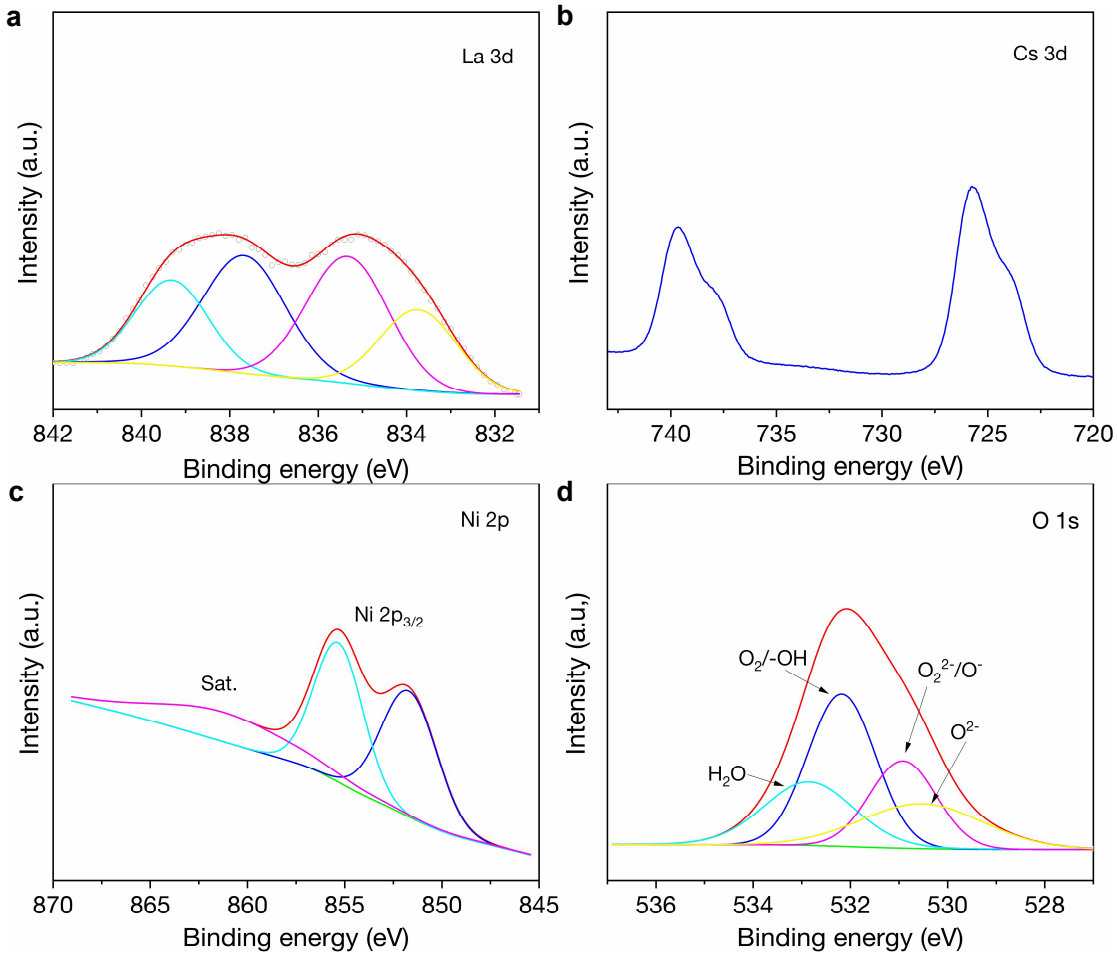

Figure S17 | XPS spectra for(a)La 3d, (b)Cs 3d, (c)Ni 2p, (d)O 1s of $Cs_{0.4}La_{0.6}NiO_3$

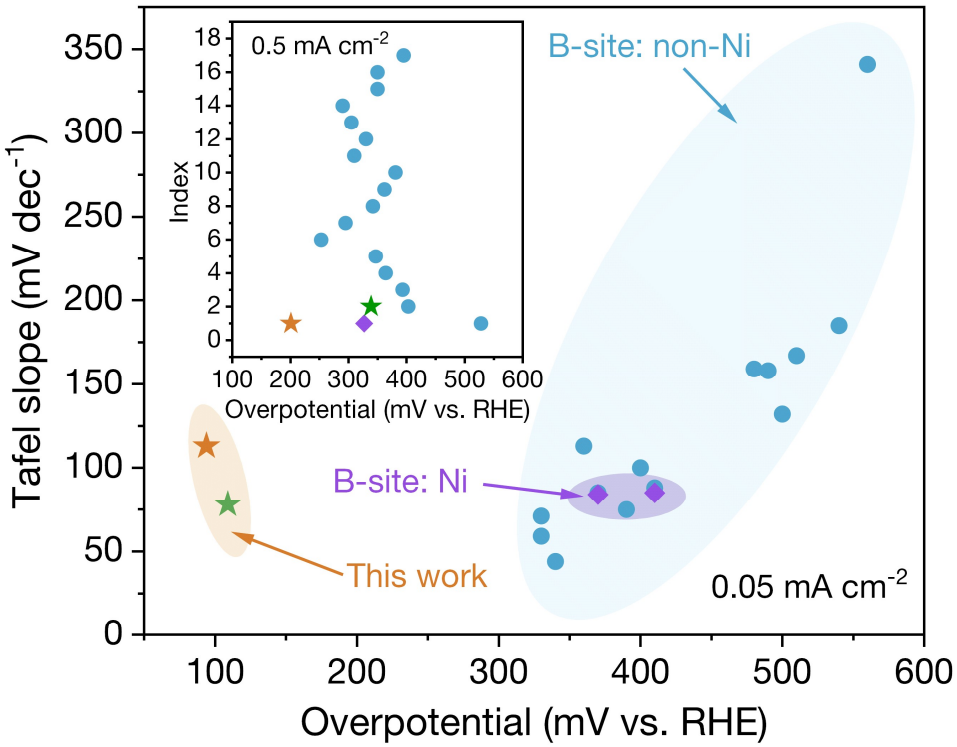

Figure S18 | Comparison of the Tafel slopes and overpotentials at 0.05mA cm^{-2} and 0.5mA cm^{-2} with perovskites in training data.

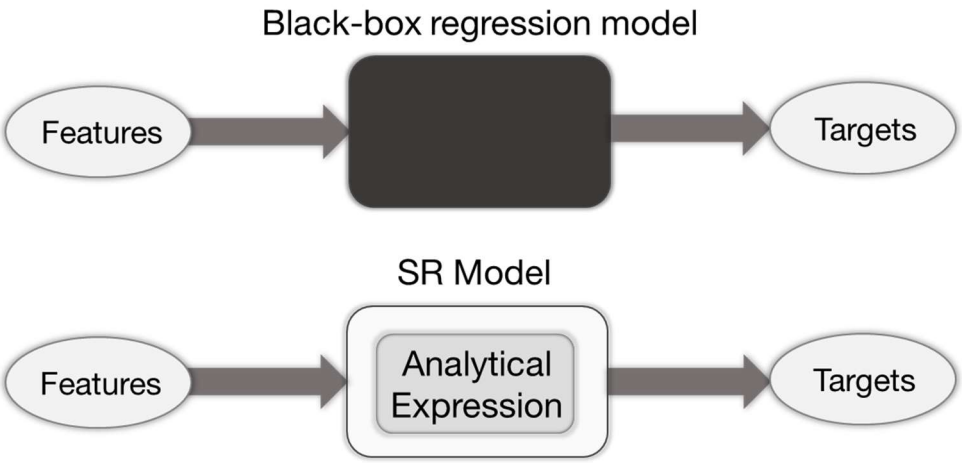

Figure S19 | Key difference between SR model and black-box regression model.

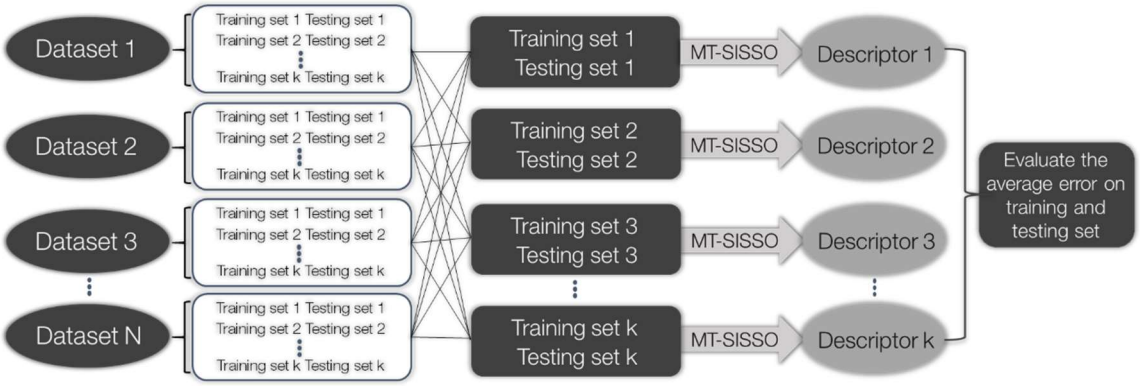

Figure S20 | Schematic diagram of *k*-fold CV process of MT-SISSO.

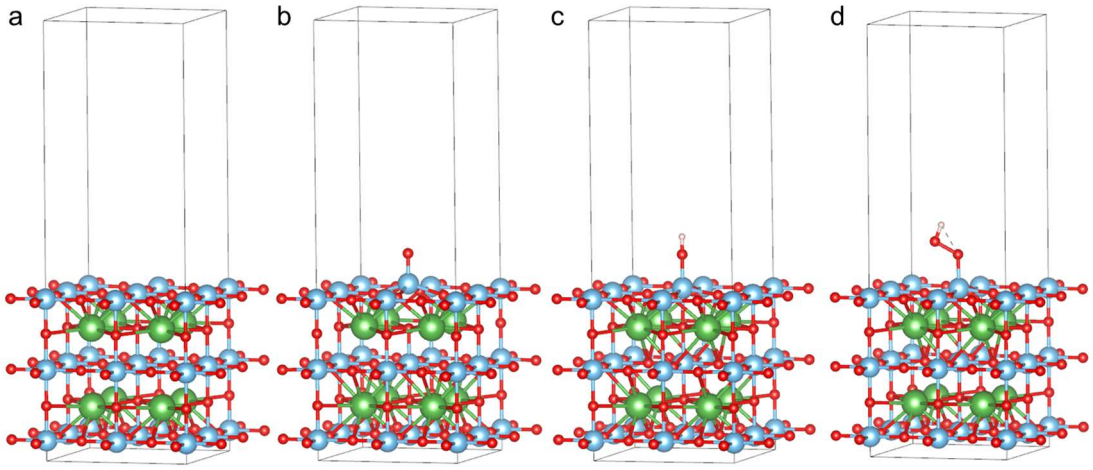

Figure S21 | An example of structures used for DFT calculations. (a) The (001) surface structures of perovskites. The adsorption structures of three intermediates (b) O, (c) OH and (d) OOH on (001) surface.

Supplementary Reference

[1] F. Nogueira, Bayesian Optimization: Open source constrained global optimization tool for Python, (2014).

[2] W. Webber, A. Moffat, J. Zobel, A similarity measure for indefinite rankings, ACM Trans. Inf. Syst. 28 (2010) 1–38.