Supporting Information for

Center-Environment Feature Model for Machine Learning Study of Spinel Oxides Based on First-Principles Computations

Yihang Li[†], Bin Xiao[†], Yuchao Tang[‡], Fu Liu[‡], Xiaomeng Wang[‡], Feinan Yan[§], and Yi Liu[†, ‡,*]

[†] Materials Genome Institute, Shanghai University, Shanghai, 200444, China

[‡] Department of Physics, College of Science, Shanghai University, Shanghai, 200444, China

[§] Department of Physics, College of Science, University of Shanghai for Science and Technology, Shanghai, 200093, China

[*]Corresponding author: yiliu@shu.edu.cn (Yi Liu)

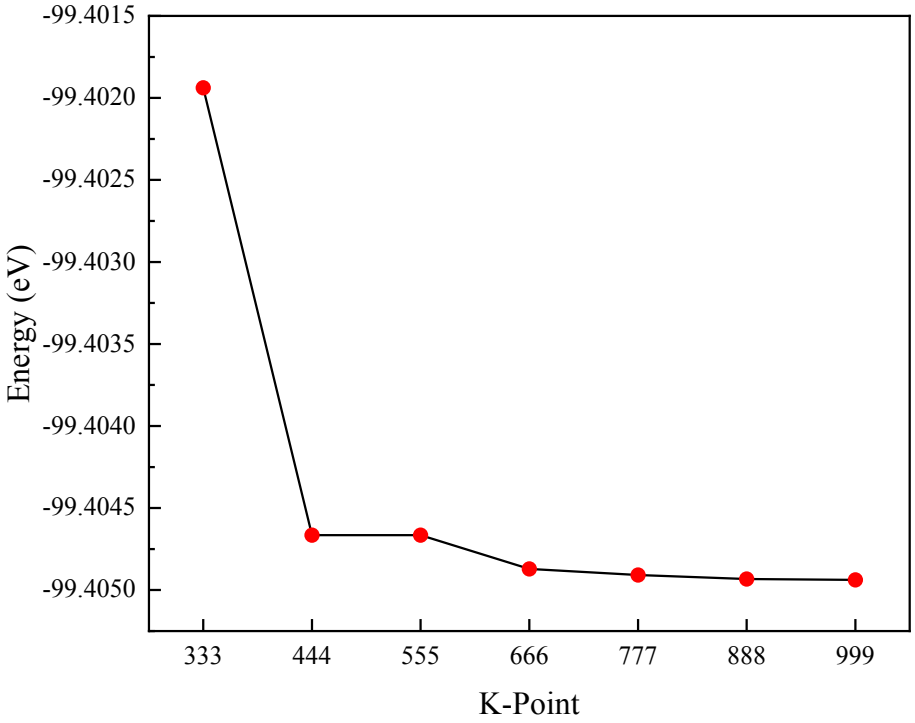

Fig. S1. According to the convergence test of k-point, integrations in the Brillouin zone were performed using 7 × 7 × 7 k-points grid.

Table S1. The GGA-PBE pseudo-potentials used for each element species in this study

Li_sv	Be	B	Na_pv	Mg	Al	Si	K_sv	Ca_ sv	Sc_sv	Ti sv
V_sv	Cr_pv	Mn_pv	Fe	Co	Ni	Cu	Zn	Ga_d	Ge_d	As
Rb_sv	Sr_sv	Y_sv	Zr_sv	Nb_sv	Tc_pv	Ru_pv	Ru_pv	Rh_pv	Pd	Ag
Cd	In_d	Sn_d	Sb	Te	Cs_sv	Ba_sv	La	Ce	Pr_3	Nd_3
Pm_3	Sm_3	Eu_2	Gd_3	Tb_3	Dy_3	Ho_3	Er_3	Tm_3	Yb_2	Lu_3
Hf_pv	Ta_pv	W_pv	Re	Os	Ir	Pt	Au	Hg	Tl_d	Pb_d
Bi_d	Ac	Th	Pa	U	Np	Pu				

Text S1. Some technical details for NuSVR

Given the training sample $D = \{(X_1, Y_1), (X_2, Y_2), ..., (X_M, Y_M)\}$, $Y_i \in R$, we hope to find a regression plane so that all the data are closest to the plane. This is called support vector regression (SVR). Suppose that we can tolerate the maximum deviation of ε between *F(x)* and *y*, that is, only when the absolute value of the difference between *F(x)* and *y* is greater than ε, then the SVR problem can be formalized as:

$$min_{w,b}\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m} l_\varepsilon(f(x_i),y_i)$$

where, *C* is the regularization constant and l_ε is the ϵ - insensitive loss function:

$$l_\varepsilon(z) = \begin{cases} 0. & |z| < \varepsilon \\ |z| - \varepsilon, & otherwise \end{cases}$$

The key codes of SVR model training in this study are as follows:

```
from sklearn.svm import NuSVR
nusvr = NuSVR()
parameters = [{'C': [0.1, 1e0, 1e1, 1e2,1e3], 'gamma': [0.001,0.01, 0.1, 0.5], 'kernel': ['rbf']}]
clf1 = GridSearchCV(nusvr, parameters, scoring='neg_mean_absolute_error', cv=5)
clf1.fit(X_train, y_train)
param = clf1.best_params_
svr1=NuSVR(C=param.get("C"),gamma=param.get("gamma"),kernel=param.get ("kernel"))
svr1.fit(X_train,y_train)
y_train_pred = svr1.predict(X_train)
y_test_pred = svr1.predict(X_test)
```

Table S2. Total energies at ground state of oxygen molecules and 73 elements in spinel oxides.

Element	**Energy**	**Ele**	**Ene**	**Ele**	**Ene**	**Ele**	**Ene**	**Ele**	**Ene**
Ac	-4.0599	Cu	-3.7281	Li	-1.9038	Pr	-4.7028	Tc	-10.3507
Ag	-2.7164	Dy	-4.5350	Lu	-4.4493	Pt	-6.1022	Te	-3.1427
Al	-3.7493	Er	-4.4981	Mg	-1.5059	Pu	-13.9806	Th	-7.4532
As	-4.6721	Eu	-10.2529	Mn	-8.9903	Rb	-0.9388	Ti	-7.8363
Au	-3.2178	Fe	-8.2385	Mo	-10.9340	Re	-12.4245	Tl	-2.2278
B	-6.7042	Ga	-2.9045	Na	-1.3095	Rh	-7.2570	Tm	-4.4443
Ba	-1.9083	Gd	-4.5821	Nb	-10.2177	Ru	-9.1229	U	-11.1377
Be	-3.7649	Ge	-4.5182	Nd	-4.7159	Sb	-4.1409	V	-8.9945
Bi	-3.8936	Hf	-9.9249	Ni	-5.4681	Sc	-6.2477	W	-12.9550
Ca	-1.9255	Hg	-0.1880	Np	-12.7260	Si	-5.4252	Y	-6.4280
Cd	-0.7476	Ho	-4.5148	Os	-11.2488	Sm	-4.6494	Yb	-1.4652
Ce	-5.9219	In	-2.5595	Pa	-9.3966	Sn	-3.8503	Zn	-1.1057
Co	-7.0181	Ir	-8.8499	Pb	-3.5646	Sr	-1.6384	Zr	-8.5193
Cr	-9.5128	K	-1.0510	Pd	-5.2205	Ta	-11.8148	O_2	-9.8550
Cs	-0.8518	La	-4.8858	Pm	-4.6878	Tb	-4.5583		

Table S3. Atomic and structural properties of elements and pure substances as elemental properties.

Property	Property
Atomic weight	Distance from core electron (Schubert)
Boiling temperature	Distance from valence electron (Schubert)
Charge nuclear effective (Clementi)	Electrochemical weight equivalent
Compression modulus	Electronegativity (Alfred-Rochow)
Density	Electronegativity (Martynov&Batsanov)
Melting temperature	Electronegativity (Pauling)
Mendeleev chemists sequence	Electronegativity absolute
Mendeleev d-t start left	Energy cohesive (Brewer)
Mendeleev d-t start right	Energy of ionization first
Mendeleev H d-t start left	Energy of ionization second
Mendeleev H d-t start right	Energy of ionization third
Mendeleev H t-d start left	Enthalpy of atomization
Oxidation state first	Enthalpy of melting
Quantum number	Enthalpy of vaporization
Radii covalent	Entropy of solid
Radii metal (Waber)	Group number
Mendeleev H t-d start right	Molar heat capacity
Mendeleev Pettifor	Moment nuclear magnetic
Mendeleev Pettifor regular	Nuclear charge effective (Slater)
Mendeleev t-d start left	Thermal neutron capture cross section
Mendeleev t-d start right	Valence electron number
Mass attenuation coefficient for MoKα	Radii pseudo-potential (Zunger)
Spectral lines n0	Magnetic resonance
Spin nuclei	Volume of atom (Villars, Daams)
Magnetic frequency of nuclei	Atomic electron scattering factor at 0.5
Periodic number start counting top right, right-left sequence	Periodic number start counting left bottom, left-right sequence
Atomic number start counting left top, left-right sequence	Atomic environment number (Villars, Daams)

Table S4. Data tables of training and testing sets whose rows represent each sample of the 5329 spinel structures and columns are the features constructed based on the CE models described above plus the target values calculated by DFT.

No.	Structure	Feature1	Feature2	Feature3	Feature4	Feature5	Feature N	Target
1	$AcAc_2O_4$	5.084	17.597	130.850	331.216	4972.601	…	-3.035
2	$AcAg_2O_4$	4.656	17.489	117.846	294.507	4638.045	…	-0.962
3	$AcAl_2O_4$	4.125	17.500	107.989	271.061	4730.198	…	-2.559
4	$AcAs_2O_4$	4.505	15.746	113.783	284.944	4185.670	…	-1.371
5	$AcAu_2O_4$	4.965	17.491	127.163	320.446	4826.720	…	-0.783
6	AcB_2O_4	4.079	14.295	105.664	266.367	5078.612	…	-0.837
7	$AcBa_2O_4$	4.766	18.181	120.800	303.913	4498.245	…	-2.332
8	$AcBe_2O_4$	4.068	17.501	105.367	265.826	4876.953	…	-1.926
9	$AcBi_2O_4$	5.029	15.786	129.040	325.771	4478.628	…	-1.622
k	…	…	…	…	…	…	…	…
5320	$ZrTi_2O_4$	2.583	17.589	55.928	125.546	5416.670	…	-2.645
5321	$ZrTl_2O_4$	3.205	17.559	73.534	172.270	4867.358	…	-1.573
5322	$ZrTm_2O_4$	3.095	17.578	70.035	161.889	4953.292	…	-3.022
5323	ZrU_2O_4	3.307	15.771	76.905	182.544	5556.518	…	-2.521
5324	ZrV_2O_4	2.601	14.451	56.367	126.796	5352.275	…	-2.242
5325	ZrW_2O_4	3.148	197.737	71.613	166.531	6144.992	…	-1.273
5326	ZrY_2O_4	2.785	17.574	61.005	137.805	5312.543	…	-3.012
5327	$ZrYb_2O_4$	3.111	17.612	70.376	163.197	4807.231	…	-3.385
5328	$ZrZn_2O_4$	2.688	17.623	58.456	131.106	4724.398	…	-2.144
5329	$ZrZr_2O_4$	2.803	17.614	61.423	138.777	5521.554	…	-2.662

Table S5. Formation energies of spinel oxides calculated by GGA and GGA + U methods

Structure	Formation Energy GGA (eV/atom)	Formation Energy GGA+U (eV/atom)	Difference	U Values (eV)	
$CoFe_2O_4$	-1.086	-1.901	-0.815	Co: 3.32	Fe: 5.3
$NiFe_2O_4$	-1.070	-1.880	-0.810	Ni: 6.2	Fe: 5.3
$MnFe_2O_4$	-1.109	-2.338	-1.229	Mn: 3.9	Fe: 5.3
MnV_2O_4	-1.962	-2.009	-0.047	Mn: 3.9	V: 3.25
$NiMn_2O_4$	-1.413	-2.080	-0.667	Ni: 6.2	Mn: 3.9
FeV_2O_4	-1.909	-2.586	-0.677	Fe: 5.3	V: 3.25
$NiCr_2O_4$	-1.612	-2.659	-1.047	Ni: 6.2	Cr: 3.7
CoV_2O_4	-1.749	-2.517	-0.768	Co: 3.32	V: 3.25
$MnCr_2O_4$	-1.759	-3.132	-1.373	Mn: 3.9	Cr: 3.7
$FeCr_2O_4$	-1.306	-3.015	-1.709	Fe: 5.3	Cr: 3.7
$NiCo_2O_4$	-0.971	-1.195	-0.224	Ni: 6.2	Co: 3.32
$MnCo_2O_4$	-1.022	-1.267	-0.245	Mn: 3.9	Co: 3.32

Text S2.

The formation energies of all compounds are adjusted according to the modified molecular energy of O_2. Data files can be obtained from this website：
https://github.com/liyihang1024/ML-CE-Spinel

To confirm further stability of the predicted spinel oxides, we calculated the energies of 465 experimental oxides that are possible decomposition products of 82 AB_2O_4 spinel oxides synthesized experimentally. Then we made comparisons between reactants and products of the following decomposition reactions.

$$AB_2O_4 \rightarrow AO+B_2O_3 \quad (1)$$

$$AB_2O_4 \rightarrow BO+ABO_3 \quad (2)$$

$$AB_2O_4 \rightarrow 2BO+AO_2 \quad (3)$$

$$AB_2O_4 \rightarrow BO_2+ABO_2 \quad (4)$$

We calculated the formation energies of some spinel with reference to the most stable binary oxides, as shown in Table S6. The formation energy is defined as follows:

$$E_f = E_{AB_2O_4} - E_{A_xO_y} - E_{B_xO_y} \quad (5)$$

where E_f and $E_{AB_2O_4}$ are the formation energy and total energy of spinel oxides, respectively. $E_{A_xO_y}$ and $E_{B_xO_y}$ represent the energies of the most stable binary oxides A_xO_y and B_xO_y, respectively. The reaction energies show that 43 out of 82 “stable” spinel oxide structures are more stable than their decomposed oxides at least following one of the above reactions. These reaction energy results confirmed more than half of the predicted “stable” structures are thermodynamically more stable than the simpler counterparts and the others are possibly metastable. It implies that the predicted “more stable” structures are more likely to be thermodynamically stable or metastable. The formation energy results gave the initial quick assessment on stability. Further confirmation on decomposition reaction energies are necessary before experiment synthesis.

Table S6. Formation energies of spinel oxides and corresponding binary oxides calculated by DFT. The structures in Bold have all positive reaction energies. The structures in Italic have at least one negative reaction energies.

Structure	AB_2O_4	AO	B_2O_3	AO+ B_2O_3 -AB_2O_4	BO	ABO_3	BO+ ABO_3 -AB_2O_4	AO_2	2BO+ AO_2 -AB_2O_4	BO_2	ABO_2	BO_2+ ABO_2 -AB_2O_4
$FeFe_2O_4$	-47.20	-13.44	-33.75	***0.02***	-13.44	-33.75	0.02	-18.44	1.88	-18.44	-26.88	1.88
$ZnFe_2O_4$	-42.95	-9.10	-33.75	***0.10***	-13.44	-27.35	2.16	-13.40	2.67	-18.44	-22.01	2.49
$CoFe_2O_4$	-45.97	-12.30	-33.75	*-0.07*	-13.44			-17.23	1.86	-18.44	-22.67	4.86
$NiFe_2O_4$	-43.27	-9.94	-33.75	*-0.41*	-13.44	-28.64	1.20	-13.82	2.57	-18.44		
$MgAl_2O_4$	-49.76	-11.97	-37.41	***0.38***	-11.11	-30.85	7.80	-16.63	10.91	-20.28		
$MgFe_2O_4$	-45.78	-11.97	-33.75	***0.07***	-13.44	-30.68	1.67	-16.63	2.28	-18.44	-24.52	2.81

$LiMn_2O_4$	-48.09		-37.84		-15.71			-14.54	2.13	-21.43	-26.56	***0.09***
$MnFe_2O_4$	-49.82	-15.71	-33.75	***0.36***	-13.44	-35.74	0.65	-21.43	1.51	-18.44	-27.97	3.40
$CuFe_2O_4$	-43.42	-10.25	-33.75	*-0.57*	-13.44			-14.21	2.34	-18.44	-24.12	0.86
$CoCo_2O_4$	-42.02	-12.30	-27.82	1.90	-12.30	-27.82	1.90	-17.23	0.19	-17.23	-24.60	***0.19***
MnV_2O_4	-50.11	-15.71	-40.18	*-5.78*	-15.83	-39.08	-4.80	-21.43	-2.98	-23.35		
$ZnCr_2O_4$	-49.55	-9.10	-39.92	***0.53***	-15.85	-30.89	2.81	-13.40	4.45	-22.32	-24.38	2.86
$CoAl_2O_4$	-50.07	-12.30	-37.41	***0.37***	-11.11	-33.57	5.39	-17.23	10.62	-20.28		
$ZnAl_2O_4$	-46.94	-9.10	-37.41	***0.43***	-11.11			-13.40	11.32	-20.28		
$NiMn_2O_4$	-46.32	-9.94	-37.84	*-1.45*	-15.71	-31.61	-1.00	-13.82	1.08	-21.43	-25.37	-0.48
$MgCr_2O_4$	-52.38	-11.97	-39.92	***0.49***	-15.85	-33.28	3.26	-16.63	4.06	-22.32	-27.49	2.58
FeV_2O_4	-53.32	-13.44	-40.18	*-0.30*	-15.83			-18.44	3.22	-23.35		
$FeAl_2O_4$	-51.07	-13.44	-37.41	***0.23***	-11.11	-35.41	4.55	-18.44	10.41	-20.28		
$CuCr_2O_4$	-50.01	-10.25	-39.92	*-0.16*	-15.85	-29.66	4.51	-14.21	4.10	-22.32	-27.41	0.28
ZnV_2O_4	-48.95	-9.10	-40.18	*-0.33*	-15.83	-32.05	1.08	-13.40	3.90	-23.35	-24.59	1.02
$CdFe_2O_4$	-41.75	-7.91	-33.75	***0.09***	-13.44	-26.54	1.78	-12.45	2.43	-18.44		
$CuMn_2O_4$	-47.79	-10.25	-37.84	*-0.30*	-15.71	-31.51	0.56	-14.21	2.15	-21.43	-26.27	0.08
$NiCr_2O_4$	-49.55	-9.94	-39.92	*-0.31*	-15.85			-13.82	4.03	-22.32		
CoV_2O_4	-52.87	-12.30	-40.18	***0.39***	-15.83	-35.25	1.80	-17.23	3.99	-23.35		
$ZnGa_2O_4$	-39.60	-9.10	-30.18	***0.31***	-10.06	-22.81	6.73	-13.40	6.08	-16.29		
$NiAl_2O_4$	-47.23	-9.94	-37.41	*-0.12*	-11.11	-29.13	6.99	-13.82	11.19	-20.28		
$MgMn_2O_4$	-48.32	-11.97	-37.84	*-1.48*	-15.71	-33.74	-1.13	-16.63	0.27	-21.43	-27.63	-0.74
LiV_2O_4	-51.09		-40.18		-15.83	-34.21	1.06	-14.54	4.90	-23.35	-27.79	*-0.04*
$SiFe_2O_4$	-50.74	-13.78	-33.75	3.21	-13.44	-37.30	***0.00***	-23.75	0.11	-18.44		
$MnCr_2O_4$	-56.21	-15.71	-39.92	***0.57***	-15.85			-21.43	3.07	-22.32		
$FeCr_2O_4$	-53.81	-13.44	-39.92	***0.45***	-15.85	-36.79	1.17	-18.44	3.67	-22.32		
MgV_2O_4	-51.78	-11.97	-40.18	*-0.37*	-15.83	-35.64	0.31	-16.63	3.50	-23.35	-26.82	1.61
$MgGa_2O_4$	-42.41	-11.97	-30.18	***0.26***	-10.06			-16.63	5.67	-16.29		
$ZnMn_2O_4$	-44.97	-9.10	-37.84	*-1.97*	-15.71	-30.60	-1.35	-13.40	0.14	-21.43	-24.26	-0.72
$CuRh_2O_4$	-43.20	-10.25	-32.87	0.09				-14.21		-19.54	-23.83	*-0.17*
$SiMg_2O_4$	-47.52	-13.78			-11.97	-35.96	*-0.40*	-23.75	-0.16	-16.63		
$SiNi_2O_4$	-43.56	-13.78	-22.46	7.32	-9.94			-23.75	*-0.06*	-13.82		
$CuAl_2O_4$	-47.21	-10.25	-37.41	*-0.44*	-11.11	-30.91	5.19	-14.21	10.78	-20.28	-26.04	0.89
$CdCr_2O_4$	-48.30	-7.91	-39.92	***0.47***	-15.85			-12.45	4.15	-22.32		
$ZnCo_2O_4$	-39.17	-9.10	-27.82	2.25	-12.30	-26.05	***0.82***	-13.40	1.17	-17.23	-20.99	0.94
$MnTi_2O_4$	-60.81	-15.71	-45.09	0.01	-17.97	-42.95	*-0.10*	-21.43	3.44	-26.92		0.29
$LiNi_2O_4$	-33.77		-22.46		-9.94	-21.46	2.37	-14.54	*-0.64*	-13.82	-19.66	2.61
$SnZn_2O_4$	-36.24	-11.53			-9.10	-27.75	*-0.61*			-13.40	-20.23	
$CoGa_2O_4$	-42.75	-12.30	-30.18	***0.27***	-10.06			-17.23	5.40	-16.29		
$MgIn_2O_4$	-40.05	-11.97	-28.27	*-0.19*	-9.72	-24.23	6.09	-16.63	3.98	-15.58		0.08
$NiCo_2O_4$	-38.55	-9.94	-27.82	0.80	-12.30	-27.20	*-0.95*	-13.82	0.13	-17.23	-21.24	
$MoAg_2O_4$	-39.13		-20.24		-7.98			-22.22	***0.95***	-12.44		
$MoNa_2O_4$	-42.42					-31.84		-22.22			-25.06	***0.30***
$LiTi_2O_4$	-57.47		-45.09		-17.97			-14.54	***6.98***	-26.92	-30.24	
$HgCr_2O_4$	-45.86	-5.84	-39.92	***0.10***	-15.85			-10.08	4.09	-22.32		

$GeCo_2O_4$	-44.11		-27.82		-12.30	-31.43	0.39	-19.38	***0.14***	-17.23		3.34
$SnMg_2O_4$	-42.76	-11.53			-11.97	-30.89	*-0.09*			-16.63	-22.79	
$MnAl_2O_4$	-53.36	-15.71	-37.41	***0.24***	-11.11	-37.39	4.86	-21.43	9.71	-20.28		
$CdIn_2O_4$	-36.41	-7.91	-28.27	***0.23***	-9.72			-12.45	4.52	-15.58		
$ZnIn_2O_4$	-37.17	-9.10	-28.27	*-0.21*	-9.72			-13.40	4.32	-15.58		
$GeNi_2O_4$	-39.78		-22.46		-9.94			-19.38	***0.53***	-13.82		
$MnIn_2O_4$	-44.03	-15.71	-28.27	***0.05***	-9.72	-32.98	1.33	-21.43	3.16	-15.58		
$CoRh_2O_4$	-45.42	-12.30	-32.87	***0.25***		-27.86		-17.23		-19.54		
CdV_2O_4	-48.42	-7.91	-40.18	***0.32***	-15.83	-31.48	1.11	-12.45	4.32	-23.35		
$CdRh_2O_4$	-41.19	-7.91	-32.87	***0.41***		-25.34		-12.45		-19.54		
$ZnRh_2O_4$	-42.14	-9.10	-32.87	***0.17***		-24.83		-13.40		-19.54		
$MnCo_2O_4$	-42.40	-15.71	-27.82	-1.13	-12.30	-34.02	*-3.92*	-21.43	-3.63	-17.23		
$MgRh_2O_4$	-44.89	-11.97	-32.87	***0.05***		-27.93		-16.63		-19.54		1.26
$CuGa_2O_4$	-39.90	-10.25	-30.18	*-0.53*	-10.06			-14.21	5.57	-16.29	-22.35	
$AlCo_2O_4$	-44.48	-11.11	-27.82	5.55	-12.30	-33.57	*-1.39*	-20.28	-0.40	-17.23		
$NiGa_2O_4$	-36.59	-9.94	-30.18	-3.53	-10.06	-33.57	*-7.05*	-13.82	2.65	-16.29		
$MnRh_2O_4$	-48.91	-15.71	-32.87	***0.32***				-21.43		-19.54		0.72
$MgTi_2O_4$	-56.98	-11.97	-45.09	-0.07	-17.97	-39.12	*-0.11*	-16.63	4.41	-26.92	-29.34	
$NiRh_2O_4$	-41.56	-9.94	-32.87	*-1.24*				-13.82		-19.54		
$CdGa_2O_4$	-38.27	-7.91	-30.18	***0.18***	-10.06			-12.45	5.71	-16.29		
AlV_2O_4	-54.51	-11.11	-40.18	3.22	-15.83	-38.72	*-0.04*	-20.28	2.57	-23.35		
WNa_2O_4	-43.75					-33.18		-22.87			-23.58	***1.67***
$MgCo_2O_4$	-42.51	-11.97	-27.82	2.73	-12.30	-29.04	***1.17***	-16.63	1.28	-17.23	-23.61	
$GeMg_2O_4$	-43.97				-11.97	-31.81	***0.20***	-19.38	0.66	-16.63		
$SiCo_2O_4$	-47.92	-13.78	-27.82	6.32	-12.30	-35.96	*-0.34*	-23.75	-0.43	-17.23		
$SnCd_2O_4$	-34.24	-11.53			-7.91	-26.94	*-0.62*			-12.45		
$PdZn_2O_4$	-32.14				-9.10			-15.88	*-1.95*	-13.40		
$AlMn_2O_4$	-52.77	-11.11	-37.84	3.82	-15.71	-37.39	*-0.33*	-20.28	1.07	-21.43		
$GeFe_2O_4$	-46.91		-33.75		-13.44	-30.85	2.63	-19.38	***0.66***	-18.44		
$AlFe_2O_4$	-48.74	-11.11	-33.75	3.88	-13.44	-35.41	*-0.11*	-20.28	1.58	-18.44		1.27
$CuCo_2O_4$	-40.18	-10.25	-27.82	2.11	-12.30			-14.21	***1.37***	-17.23	-21.68	
$CdAl_2O_4$	-45.24	-7.91	-37.41	*-0.08*	-11.11			-12.45	10.57	-20.28		

Table S7. Test Set R^2 and RMSE of machine learning results for each model of total energy

R^2	1	2	3	4	5	6	7	8	9	10
CE-1	0.153	0.959	0.961	0.960	0.959	0.959	0.959	0.960	0.961	0.958
CE-2	0.993	0.995	0.996	0.995	0.996	0.996	0.995	0.996	0.996	0.996
CE-3	0.782	0.786	0.780	0.984	0.983	0.983	0.981	0.978	0.983	0.982
CE-4	0.843	0.995	0.995	0.992	0.992	0.992	0.992	0.993	0.992	0.992
CE-5	0.842	0.991	0.992	0.977	0.974	0.968	0.982	0.988	0.974	0.979
CE-6	0.150	0.994	0.994	0.994	0.994	0.993	0.995	0.995	0.993	0.993
CE-7	0.785	0.783	0.784	0.995	0.994	0.995	0.994	0.992	0.994	0.993
CE-8	0.994	0.996	0.996	0.995	0.995	0.996	0.996	0.996	0.996	0.996
RMSE	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
CE-1	1.315	0.287	0.283	0.284	0.286	0.288	0.287	0.287	0.282	0.291
CE-2	0.116	0.097	0.090	0.096	0.093	0.094	0.097	0.093	0.090	0.093
CE-3	0.666	0.662	0.667	0.177	0.185	0.189	0.196	0.209	0.182	0.190
CE-4	0.562	0.101	0.101	0.127	0.130	0.125	0.124	0.115	0.125	0.126
CE-5	0.565	0.136	0.131	0.215	0.227	0.253	0.191	0.157	0.228	0.208
CE-6	1.310	0.109	0.109	0.110	0.106	0.117	0.104	0.104	0.119	0.117
CE-7	0.664	0.661	0.659	0.103	0.106	0.103	0.106	0.125	0.108	0.120
CE-8	0.108	0.089	0.089	0.097	0.097	0.095	0.093	0.090	0.096	0.094

Table S8. Test Set R^2 and RMSE of machine learning results for each model of band gap.

R^2	1	2	3	4	5	6	7	8	9	10
CE-1	0.118	0.567	0.700	0.672	0.689	0.690	0.703	0.709	0.679	0.679
CE-2	0.802	0.783	0.803	0.811	0.813	0.783	0.784	0.810	0.795	0.806
CE-3	0.226	0.250	0.444	0.667	0.640	0.617	0.625	0.544	0.666	0.644
CE-4	0.628	0.786	0.763	0.785	0.791	0.776	0.798	0.753	0.793	0.773
CE-5	0.723	0.722	0.761	0.768	0.732	0.735	0.740	0.747	0.738	0.713
CE-6	0.133	0.553	0.802	0.809	0.794	0.806	0.814	0.780	0.798	0.807
CE-7	0.183	0.267	0.527	0.795	0.773	0.799	0.806	0.740	0.751	0.719
CE-8	0.626	0.770	0.775	0.795	0.779	0.788	0.806	0.782	0.805	0.820
RMSE	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
CE-1	1.223	0.820	0.706	0.742	0.688	0.721	0.710	0.699	0.733	0.729
CE-2	0.584	0.595	0.569	0.535	0.567	0.600	0.589	0.566	0.590	0.553
CE-3	1.134	1.135	0.983	0.714	0.767	0.776	0.768	0.852	0.742	0.754
CE-4	0.788	0.604	0.621	0.592	0.588	0.596	0.574	0.641	0.583	0.609
CE-5	0.665	0.675	0.638	0.624	0.647	0.651	0.667	0.639	0.657	0.691
CE-6	1.177	0.844	0.576	0.576	0.595	0.565	0.540	0.588	0.580	0.560
CE-7	1.162	1.116	0.886	0.573	0.610	0.569	0.567	0.662	0.642	0.666
CE-8	0.788	0.609	0.609	0.574	0.605	0.588	0.576	0.585	0.575	0.546

Table S9. Test Set R2 and RMSE of machine learning results for each model of lattice constant.

R^2	1	2	3	4	5	6	7	8	9	10
CE-1	0.160	0.924	0.925	0.923	0.920	0.921	0.927	0.923	0.923	0.921
CE-2	0.961	0.963	0.963	0.963	0.962	0.962	0.962	0.962	0.959	0.962
CE-3	0.757	0.755	0.754	0.955	0.955	0.953	0.954	0.953	0.954	0.951
CE-4	0.827	0.963	0.964	0.961	0.963	0.961	0.964	0.963	0.962	0.963
CE-5	0.823	0.961	0.957	0.939	0.934	0.928	0.948	0.955	0.933	0.944
CE-6	0.158	0.962	0.960	0.960	0.962	0.962	0.960	0.960	0.961	0.960
CE-7	0.759	0.755	0.758	0.963	0.962	0.964	0.964	0.961	0.962	0.962
CE-8	0.960	0.963	0.963	0.963	0.960	0.962	0.961	0.962	0.963	0.963
RMSE	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
CE-1	0.356	0.107	0.105	0.107	0.109	0.108	0.104	0.107	0.106	0.109
CE-2	0.077	0.075	0.075	0.074	0.075	0.075	0.075	0.075	0.078	0.074
CE-3	0.191	0.191	0.191	0.082	0.082	0.084	0.083	0.083	0.082	0.084
CE-4	0.160	0.074	0.073	0.076	0.075	0.077	0.072	0.075	0.075	0.074
CE-5	0.163	0.076	0.080	0.096	0.099	0.103	0.087	0.081	0.100	0.091
CE-6	0.355	0.075	0.076	0.077	0.075	0.075	0.077	0.076	0.075	0.077
CE-7	0.190	0.190	0.190	0.074	0.075	0.073	0.074	0.075	0.075	0.075
CE-8	0.077	0.074	0.074	0.073	0.077	0.075	0.076	0.075	0.074	0.074

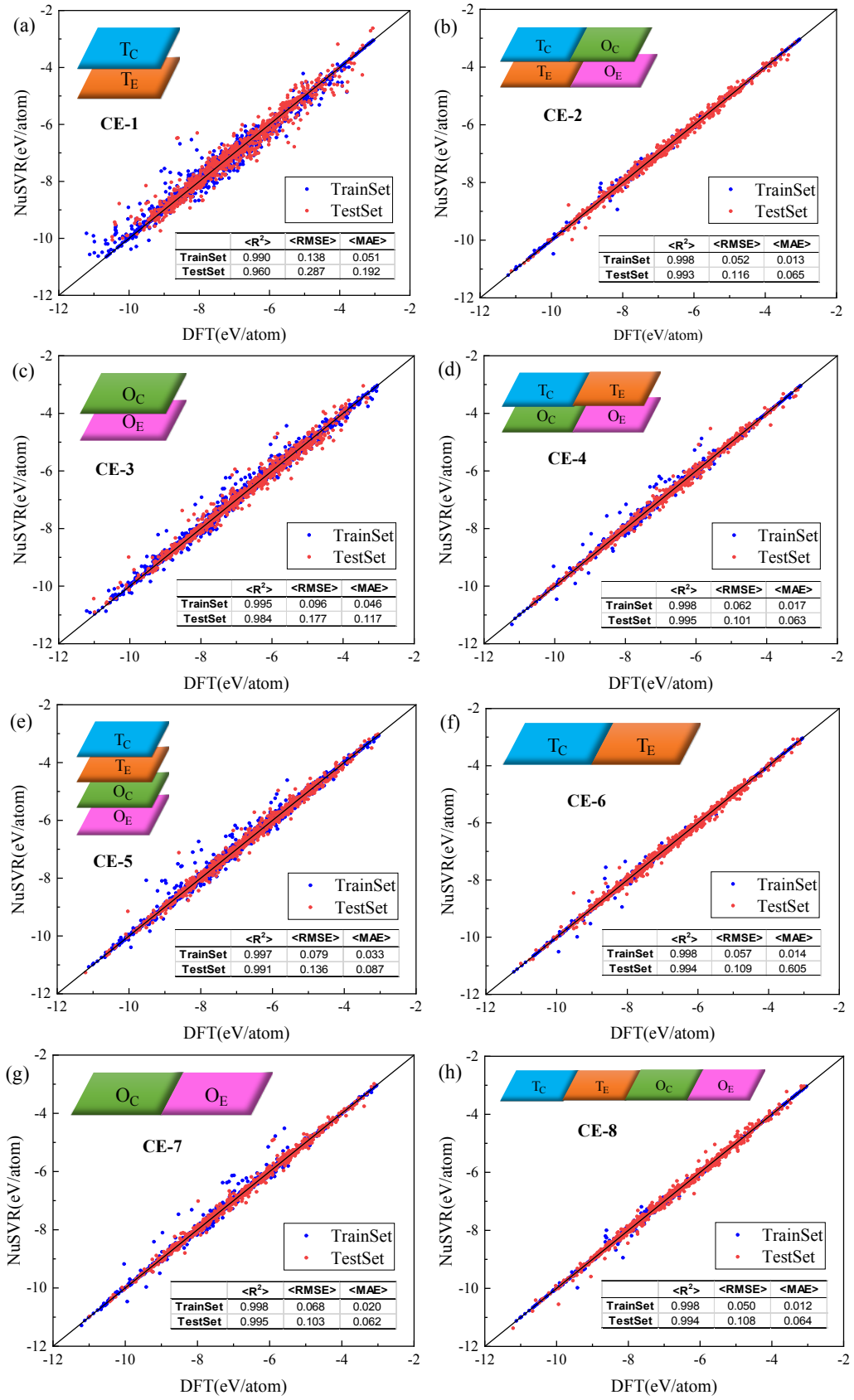

Figure S2. Comparison between the DFT calculations and machine learning predictions by SVR_rbf algorithm using CE-2 to CE-8 CE feature models at the nearest-neighbors, respectively, for the [(a) to (h)] total energy (eV/ atom).

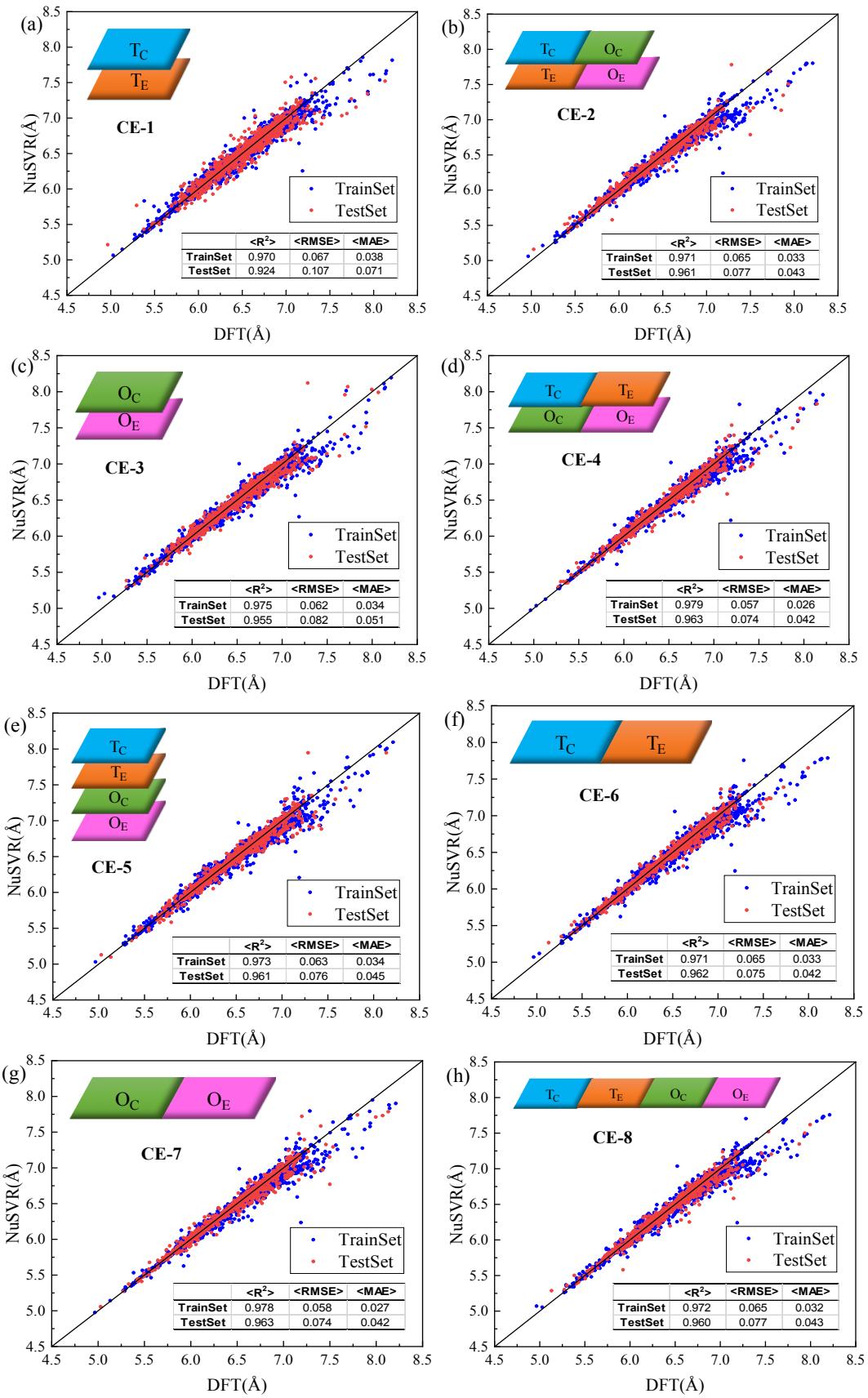

Figure S3. Comparison between the DFT calculations and machine learning predictions by SVR_rbf algorithm using CE-2 to CE-8 CE feature models at the nearest-neighbors, respectively, for the [(a) to (h)] lattice constant (Å).

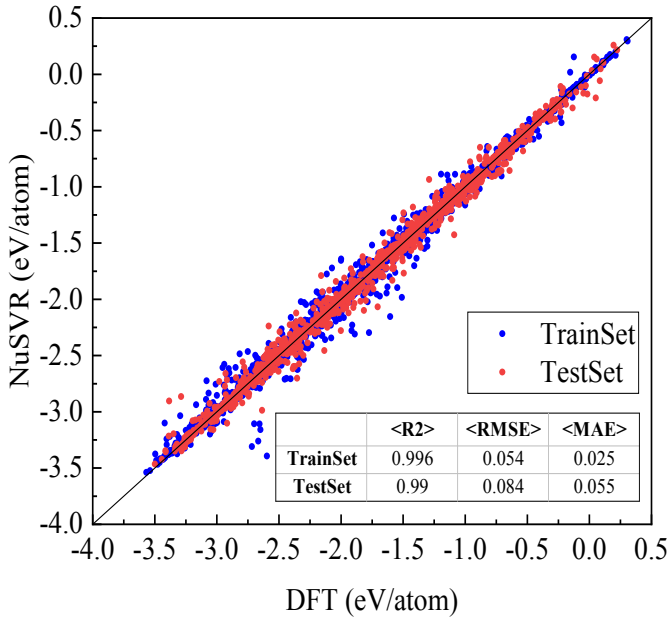

Figure S4. ML prediction results of formation energies (eV/atom) by composition models