Supplementary Material

Machine Learning for Prediction of CO_2/N_2/H_2O Selective Adsorption and Separation in Metal-Zeolites

Yating Gu, Yuming Gu[*], Qiantu Tao, Xinzhu Wang, Qin Zhu, Jing Ma[*]

Key Laboratory of Mesoscopic Chemistry of Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing, 210023, China.

Correspondence to: Dr. Yuming Gu, Department, Key Laboratory of Mesoscopic Chemistry of Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Xianlin Avenue, Nanjing, 210023, Jiangsu, China. E-mail: guyuming19930701@163.com. Prof. Jing Ma, Department, Key Laboratory of Mesoscopic Chemistry of Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Xianlin Avenue, Nanjing, 210023, Jiangsu, China. E-mail: majing@nju.edu.cn

List of Contents

S1. Structural Model

The physical properties of CO_2, N_2 and H_2O are listed in Supplementary Table 1.

Supplementary Table 1. Physical properties of CO_2, N_2 and H_2O.

Physical properties	CO_2	N_2	H_2O
Kinetic diameter (Å)	3.3	3.6	2.7
Polarizability ($\times 10^{-25}$ cm^3)	26.3	17.7	14.8
Dipole moment ($\times 10^{18}$ esu cm)	0	0	1.87
Quadrupole moment ($\times 10^{-26}$ esy cm^2)	4.30	1.52	–

In this study, five kinds of topological structures (CHA, MFI, MOR, FER and FAU zeolites) were investigated for the CO_2 adsorption and separation. The supercell models for CHA (2 × 2 × 2), MOR (1 × 1 × 2), and FER (1 × 1 × 2) were used in calculation.[1] To take the influence of the silicon-aluminum ratio (Si/Al) into consideration, models were established by replacing one/two/three/four silicon atoms with aluminum atoms, named as MFI-95, MFI-47, MFI-31, MFI-23, MOR-47, FER-35, CHA-47, and FAU-95, respectively.
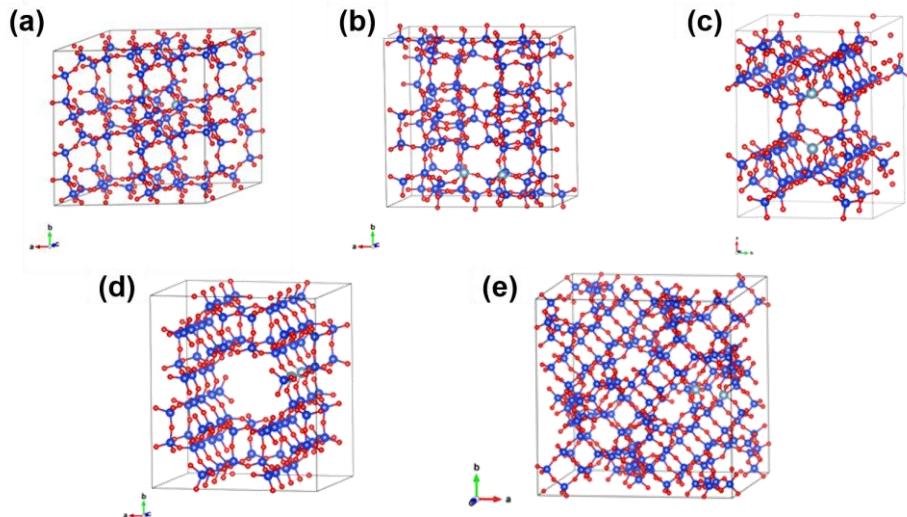

Supplementary Figure 1. Structures of (a) CHA, (b) MFI, (c) MOR, (d) FER and (e) FAU zeolites.
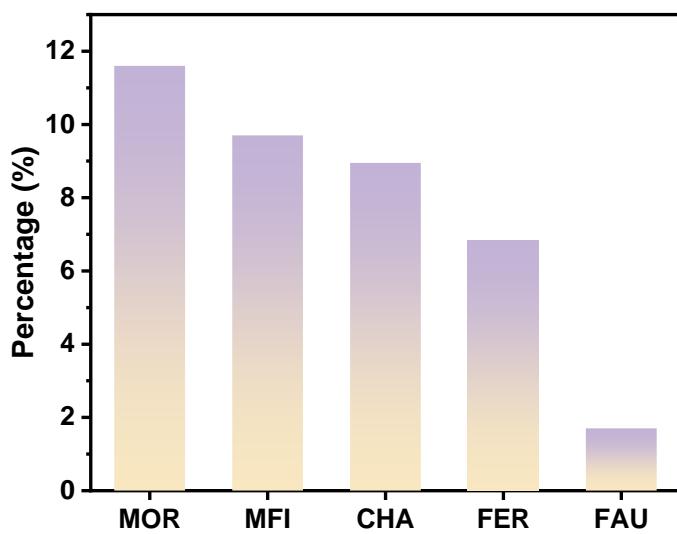
Supplementary Figure 2. The proportion of five types of topological structures (MOR, MFI, CHA, FER, FAU) in the literatures.

The number of IZA structures has reached 264. There have been 197,763 reports related to zeolite from all databases on Web of science until August 14th, 2023. Moreover, the five kinds of topological structures have been studied widely according to literature researches (Supplementary Figure 2). Among them, researches about MOR take up 11.7%, and followed by MFI, accounting for 9.8%. CHA, FER and FAU have the proportion of 9.0%, 6.9% and 1.8%, respectively.
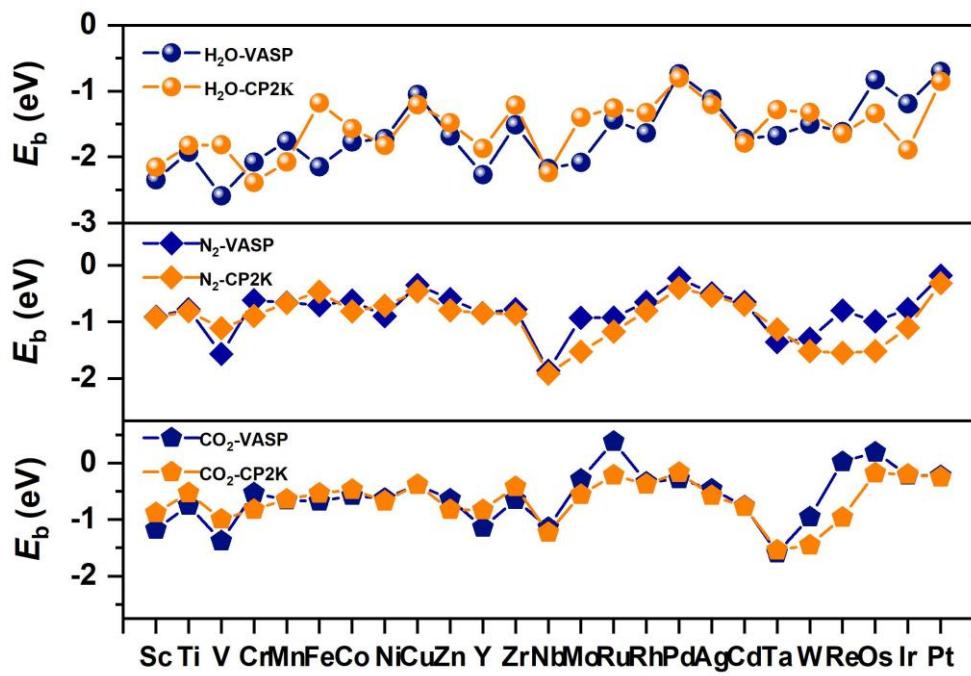
Supplementary Figure 3. The computational results of MFI-31 comparison between VASP and CP2K.
In order to validate the computational results obtained by CP2K, MFI-31 zeolites were also performed using Vienna Ab initio Simulation Package (VASP), which gave similar results to CP2K.
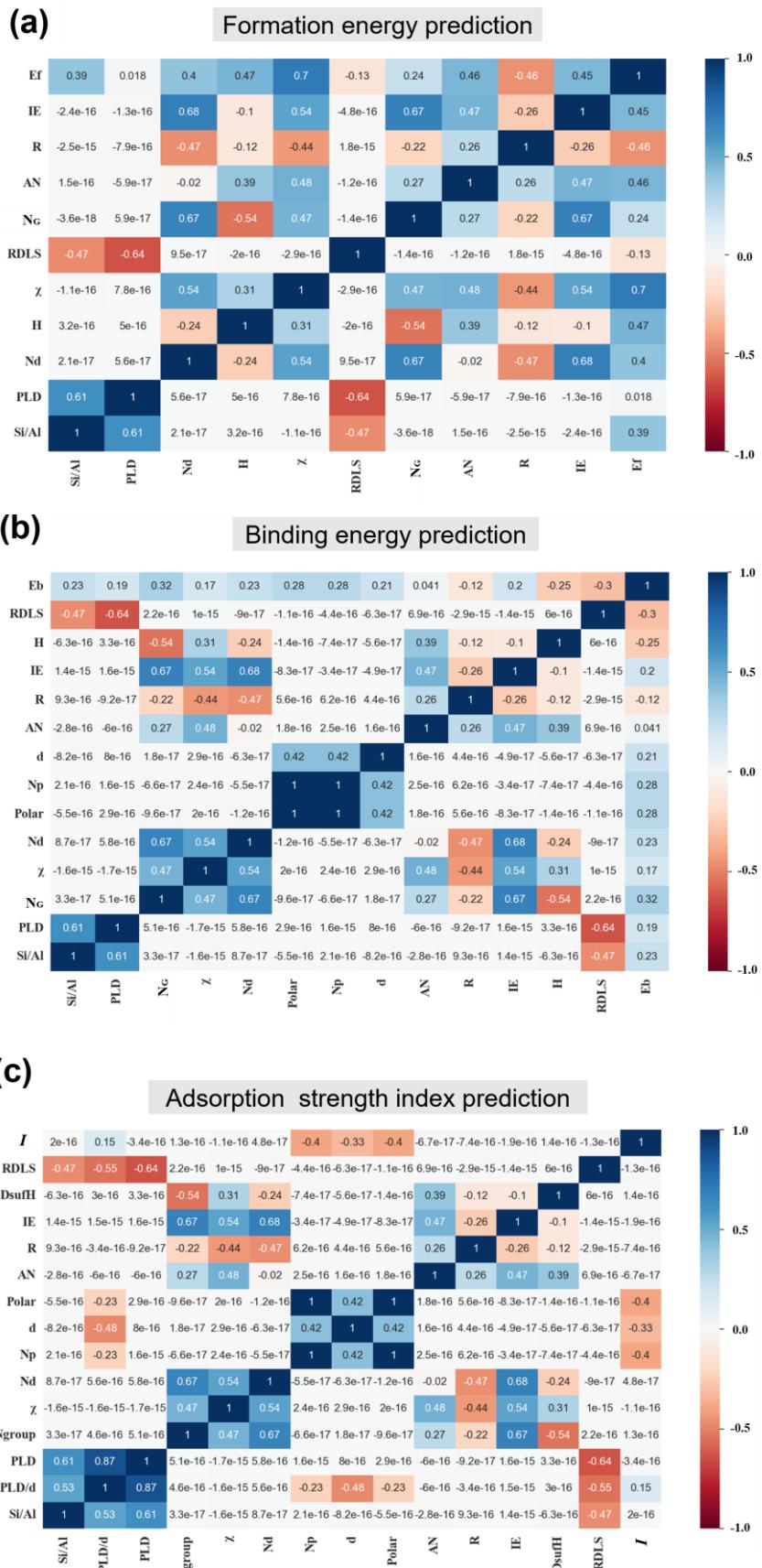
Supplementary Figure 4. Pearson correlation coefficient matrix of descriptors for predicting (a) formation energy, (b) binding energy, (c) adsorption strength index.

S2. Machine Learning

Machine learning methods, ten algorithms, namely, least absolute shrinkage and selection operator (LASSO),[2] support vector regression (SVR),[3] gradient boosting regression (GBR), extreme gradient boosting regression (XGBoost), decisiontree,[4] k-nearest neighbor (kNN),[5] linear ridge,[6] multiple Linear regression (MLR),[7] artificial neural network (ANN),[8] and ExtraTrees[9], were chosen to compare the performance of the prediction models. The LASSO is a penalized likelihood approach. The SVR algorithm applies the squared ϵ-insensitive loss function, which makes the optimization problem strictly convex and obtains a clear solution. The GBR and XGBoost are ensemble boosting learning models. [10] The ensemble learning algorithm can train multiple constituent learning algorithms at onetime to obtain better prediction compared with a single-constituent learner. Three evaluation indexes, mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2) are used to evaluate the predictive capacity of ML models. ML trainings were implemented by the aid of scikit-learn package. [11] ExtreTrees, GBR and XGBoost showed the good performance on the formation energy, binding energy and adsorption strength index.
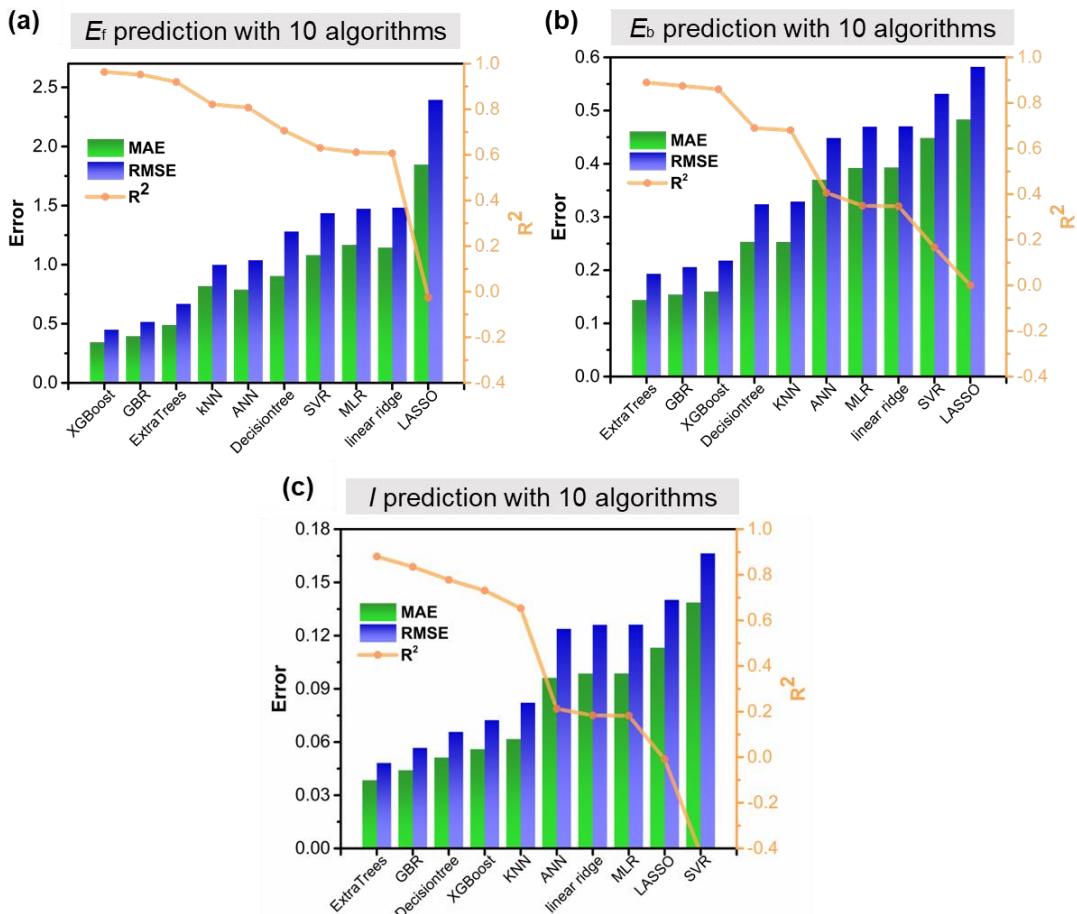
Supplementary Figure 5. Comparison of the performance with 10 algorithms of the FL in (a) formation energy, (b) binding energy, (c) adsorption strength index.
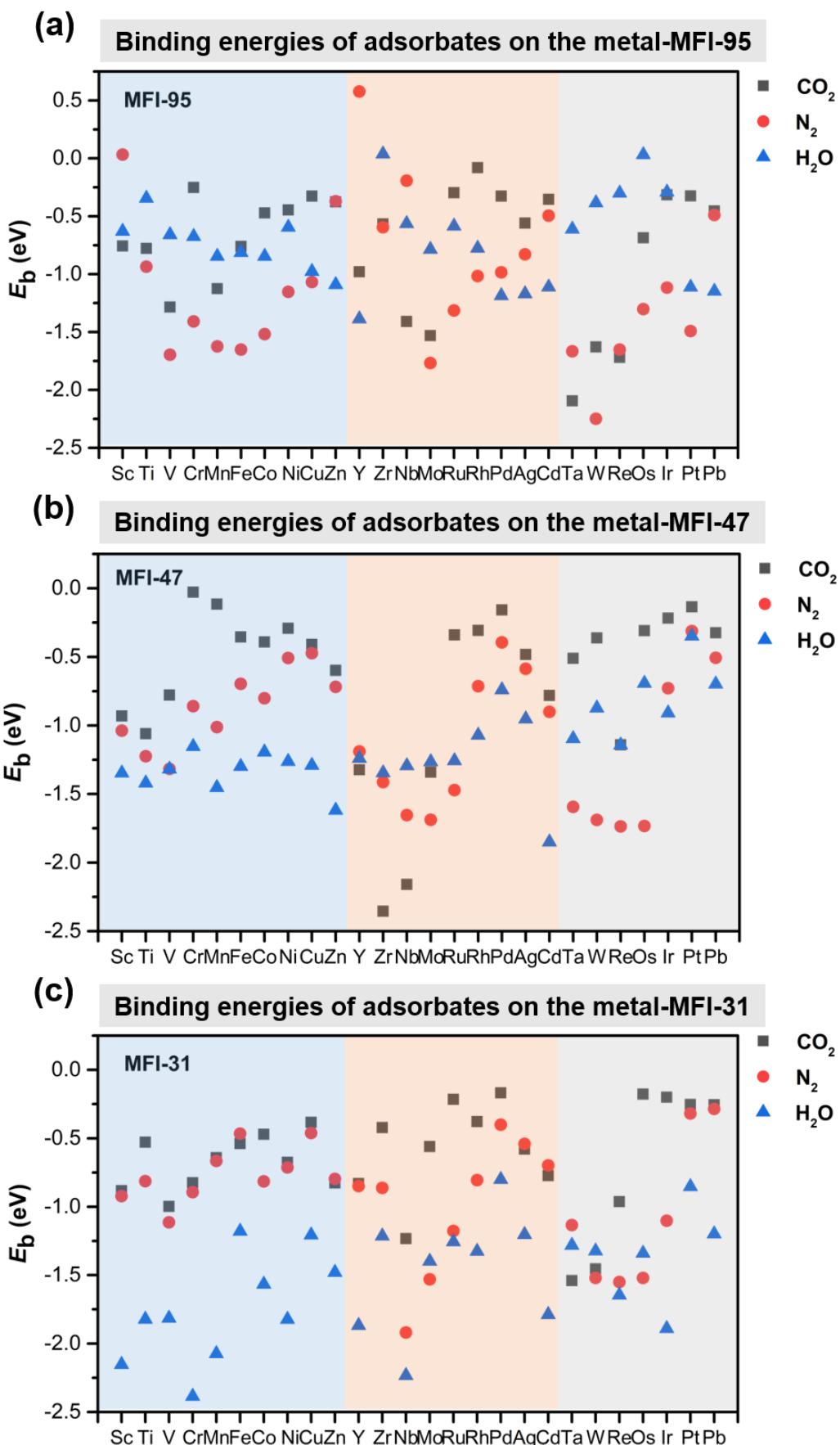
Supplementary Figure 6. The calculated binding energies of adsorbates on the metal-zeolites (a) MFI-95, (b) MFI-47, (c) MFI-31.

(a) Binding energies of adsorbates on the metal-MFI-23

(b) Binding energies of adsorbates on the metal-MOR-47

(c) Binding energies of adsorbates on the metal-FER-35
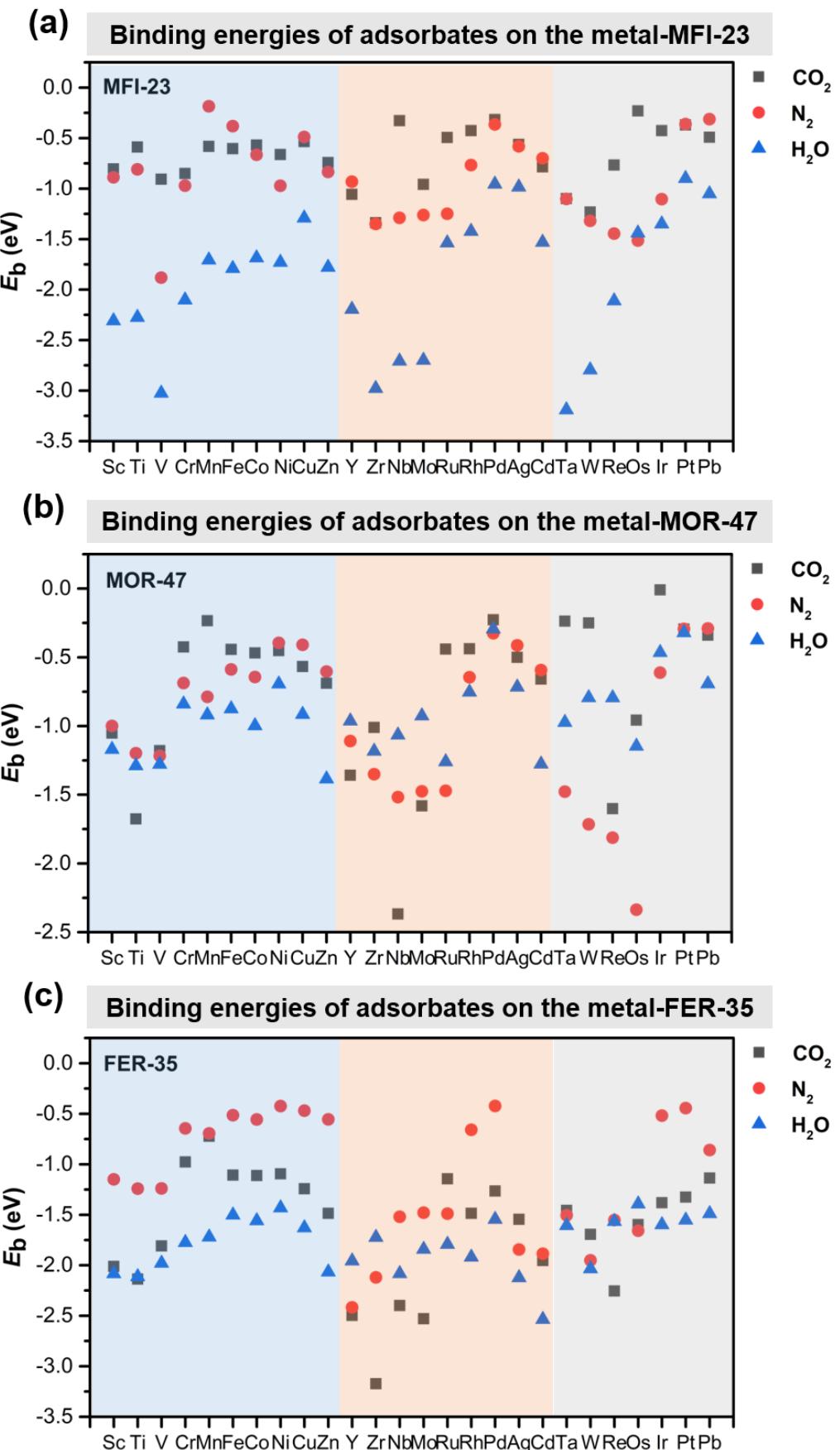
Supplementary Figure 7. The calculated binding energies of adsorbates on the metal-zeolites (a) MFI-23, (b) MOR-47, (c) FER-35.

(a)

Binding energies of adsorbates on the metal-CHA-47

(b)

Binding energies of adsorbates on the metal-FAU-95
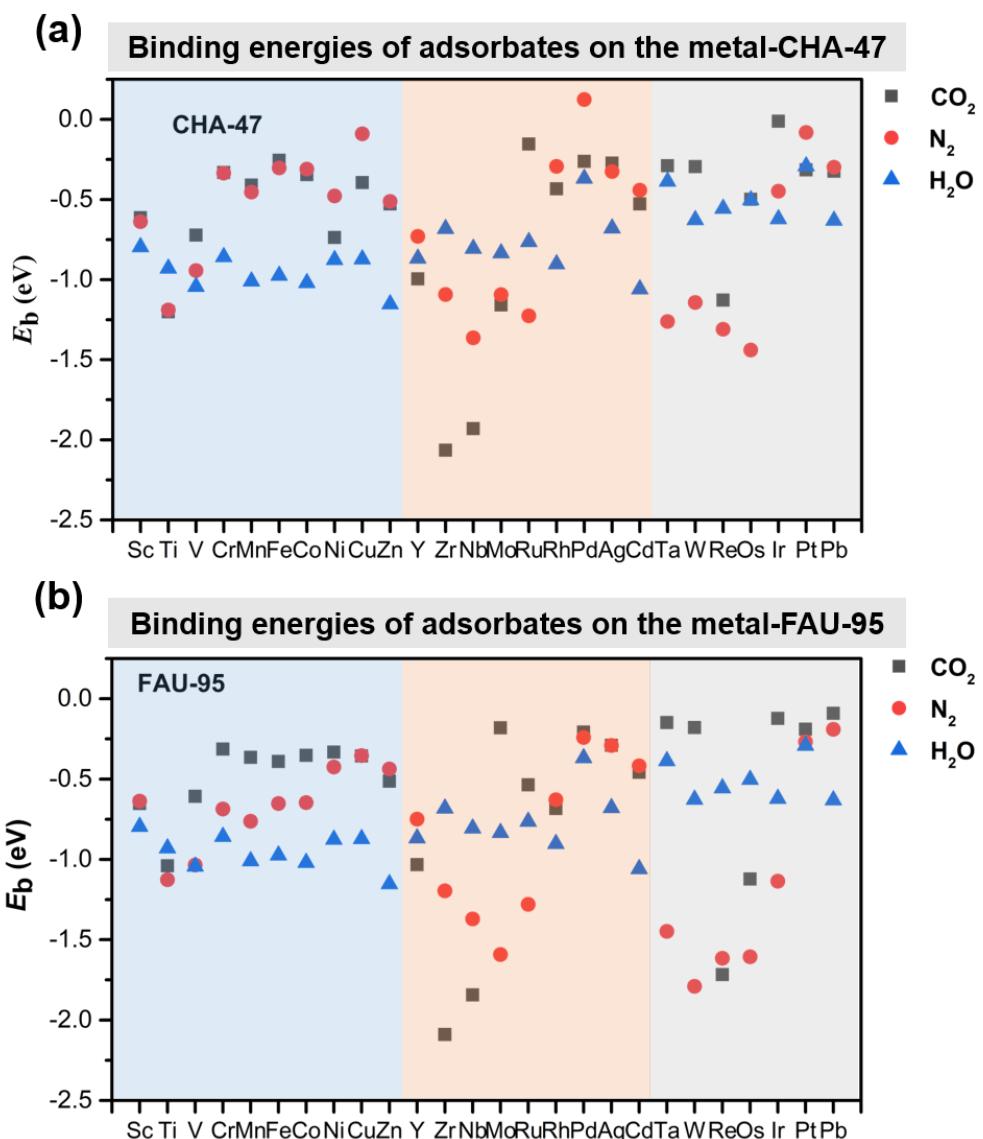
Supplementary Figure 8. The calculated binding energies of adsorbates on the metal-zeolites (a) CHA-47, (b) FAU-95.

The binding energies of 5 kinds of zeolites with embedded metals were calculated to study the adsorption strength of adsorbates. The embedded metal atoms are involved in $3d$, $4d$, and $5d$ transition metal. As shown in Supplementary Figures 6-8, the Mo, Zr-zeolite preferred to adsorb the CO_2 in comparison of the N_2 and H_2O. Zeolites with $4d$ transition metal embedded showed stronger binding ability than those of $3d$ and $5d$ transition metals.

S3. Feature Selection and the Dataset

To find out the good materials to separate the CO_2 from the mixed gases (N_2/H_2O) with the similar kinetic diameters, there are 15 features chosen to build the machine learning models to predict the formation energy, binding energy and adsorption strength index. To describe the influence of metal atoms, the metal atomic electronegativity (χ), first ionization energy of metal atoms (IE), the number of d electrons (Nd) and metal atomic radius (R) were chosen. The more electronegative the metal is, the easier the metal-zeolite are of attracting electron. The first ionization energy of metal atom is the energy required for a gaseous atom in its ground state to lose an electron in its outermost shell. It is easier to lose an electron when the atom has a lower first ionization. Nd is the number of d orbitals electrons of metal atoms in the valence shell. The refinement distance least squares (R_{DLS}) was considered to describe the geometry distortion of zeolite framework. The geometry distortion may play a part in adsorption process. Furthermore, changing the Si/Al ratio of the zeolites can also tune the electrostatic environment. The adsorption and diffusion properties are related to the channel architecture of zeolites. The pore limiting diameter (PLD) of zeolite was applied to describe the porosity of zeolites. All of the features are listed in the Supplementary Table 2 and the data are listed in the Supplementary Table 3.

Supplementary Table 2. List of features used in prediction of the formation energy, binding energy and adsorption strength index.

abbreviation	name	unit
PLD	Pore limiting diameter of zeolite	Å
Si/Al	Number ratio of silicon to aluminum	—
R_{DLS}	Refinement distance least-squares	—
Polar	polarizability of adsorbate	cm^3
Np	Valence electron number of the adsorbate	—
d	Kinetic diameter of the adsorbate	Å
N_G	The number of group in the periodic table of metal	—
χ	Electronegativity	—
IE	The first ionization energy of metal atom	kJ mol^{-1}
Nd	d shell electron number of metal atom	—
AN	Atomic number of metal	—
R	Metal atomic radius	pm
$\Delta_{fus}H$	Enthalpy of fusion of metal	kJ mol^{-1}
PLD/d	Ratio of pore limiting diameter of zeolite to kinetic diameter of adsorbate	—

Supplementary Table 3. Part of data (20%) in dataset of machine learning to predict the binding energy and adsorption strength index.

System	Si/Al	PLD/d	PLD (Å)	N_G	Polar ($\times 10^{-25}$ cm^3)	Nd	Np	$\Delta_{fus}H$ (kJ mol^{-1})
CO_2@Sc-MFI-95	95	1.927273	6.36	3	26.3	1	16	16.11
CO_2@Ti-MFI-95	95	1.927273	6.36	4	26.3	2	16	18.6
CO_2@V-MFI-95	95	1.927273	6.36	5	26.3	3	16	20.8
CO_2@Cr-MFI-95	95	1.927273	6.36	6	26.3	5	16	20
CO_2@Mn-MFI-95	95	1.927273	6.36	7	26.3	5	16	14.64
CO_2@Fe-MFI-95	95	1.927273	6.36	8	26.3	6	16	13.8
CO_2@Co-MFI-95	95	1.927273	6.36	9	26.3	7	16	16.19
CO_2@Ni-MFI-95	95	1.927273	6.36	10	26.3	8	16	17.2
CO_2@Cu-MFI-95	95	1.927273	6.36	11	26.3	10	16	13.14
CO_2@Zn-MFI-95	95	1.927273	6.36	12	26.3	10	16	7.38
CO_2@Y-MFI-95	95	1.927273	6.36	3	26.3	1	16	17.5
CO_2@Zr-MFI-95	95	1.927273	6.36	4	26.3	2	16	21
CO_2@Nb-MFI-95	95	1.927273	6.36	5	26.3	4	16	26.9
CO_2@Mo-MFI-95	95	1.927273	6.36	6	26.3	5	16	36
CO_2@Ru-MFI-95	95	1.927273	6.36	8	26.3	7	16	25.5
CO_2@Rh-MFI-95	95	1.927273	6.36	9	26.3	8	16	21.8
CO_2@Pd-MFI-95	95	1.927273	6.36	10	26.3	10	16	16.74
CO_2@Ag-MFI-95	95	1.927273	6.36	11	26.3	10	16	11.3
CO_2@Cd-MFI-95	95	1.927273	6.36	12	26.3	10	16	6.07
CO_2@Ta-MFI-95	95	1.927273	6.36	5	26.3	3	16	36
CO_2@W-MFI-95	95	1.927273	6.36	6	26.3	4	16	35

(*To be continued*)

System	Si/Al	PLD/d	PLD (Å)	N_G	Polar ($\times 10^{-25}$ cm^3)	Nd	Np	$\Delta_{fus}H$ (kJ mol^{-1})
CO_2@Os-MFI-95	95	1.927273	6.36	8	26.3	6	16	29.29
CO_2@Ir-MFI-95	95	1.927273	6.36	9	26.3	7	16	26.36
CO_2@Pt-MFI-95	95	1.927273	6.36	10	26.3	9	16	19.66
CO_2@Pb-MFI-95	95	1.927273	6.36	14	26.3	0	16	4.77
N_2@Sc-MFI-95	95	1.766667	6.36	3	17.7	1	10	16.11
N_2@Ti-MFI-95	95	1.766667	6.36	4	17.7	2	10	18.6
N_2@V-MFI-95	95	1.766667	6.36	5	17.7	3	10	20.8
N_2@Cr-MFI-95	95	1.766667	6.36	6	17.7	5	10	20
N_2@Mn-MFI-95	95	1.766667	6.36	7	17.7	5	10	14.64
N_2@Fe-MFI-95	95	1.766667	6.36	8	17.7	6	10	13.8
N_2@Co-MFI-95	95	1.766667	6.36	9	17.7	7	10	16.19
N_2@Ni-MFI-95	95	1.766667	6.36	10	17.7	8	10	17.2
N_2@Cu-MFI-95	95	1.766667	6.36	11	17.7	10	10	13.14
N_2@Zn-MFI-95	95	1.766667	6.36	12	17.7	10	10	7.38
N_2@Y-MFI-95	95	1.766667	6.36	3	17.7	1	10	17.5
N_2@Zr-MFI-95	95	1.766667	6.36	4	17.7	2	10	21
N_2@Nb-MFI-95	95	1.766667	6.36	5	17.7	4	10	26.9
N_2@Mo-MFI-95	95	1.766667	6.36	6	17.7	5	10	36
N_2@Ru-MFI-95	95	1.766667	6.36	8	17.7	7	10	25.5
N_2@Rh-MFI-95	95	1.766667	6.36	9	17.7	8	10	21.8
N_2@Pd-MFI-95	95	1.766667	6.36	10	17.7	10	10	16.74

(*To be continued*)

System	Si/Al	PLD/d	PLD (Å)	N_G	Polar ($\times 10^{-25}$ cm^3)	Nd	Np	$\Delta_{fus}H$ (kJ mol^{-1})
N$_2$@Cd-MFI-95	95	1.766667	6.36	12	17.7	10	10	6.07
N$_2$@Ta-MFI1-95	95	1.766667	6.36	5	17.7	3	10	36
N$_2$@W-MFI-95	95	1.766667	6.36	6	17.7	4	10	35
N$_2$@Re-MFI-95	95	1.766667	6.36	7	17.7	5	10	33.05
N$_2$@Os-MFI-95	95	1.766667	6.36	8	17.7	6	10	29.29
N$_2$@Ir-MFI-95	95	1.766667	6.36	9	17.7	7	10	26.36
N$_2$@Pt-MFI-95	95	1.766667	6.36	10	17.7	9	10	19.66
N$_2$@Pb-MFI-95	95	1.766667	6.36	14	17.7	0	10	4.77
H$_2$O@Sc-MFI-95	95	2.355556	6.36	3	14.8	1	8	16.11
H$_2$O@Ti-MFI-95	95	2.355556	6.36	4	14.8	2	8	18.6
H$_2$O@V-MFI-95	95	2.355556	6.36	5	14.8	3	8	20.8
H$_2$O@Cr-MFI-95	95	2.355556	6.36	6	14.8	5	8	20
H$_2$O@Mn-MFI-95	95	2.355556	6.36	7	14.8	5	8	14.64
H$_2$O@Fe-MFI-95	95	2.355556	6.36	8	14.8	6	8	13.8
H$_2$O@Co-MFI-95	95	2.355556	6.36	9	14.8	7	8	16.19
H$_2$O@Ni-MFI-95	95	2.355556	6.36	10	14.8	8	8	17.2
H$_2$O@Cu-MFI-95	95	2.355556	6.36	11	14.8	10	8	13.14
H$_2$O@Zn-MFI-95	95	2.355556	6.36	12	14.8	10	8	7.38
H$_2$O@Y-MFI-95	95	2.355556	6.36	3	14.8	1	8	17.5
H$_2$O@Zr-MFI-95	95	2.355556	6.36	4	14.8	2	8	21
H$_2$O@Nb-MFI-95	95	2.355556	6.36	5	14.8	4	8	26.9

(*To be continued*)

System	Si/Al	PLD/d	PLD (Å)	N_G	Polar ($\times 10^{-25}$ cm^3)	Nd	Np	$\Delta_{fus}H$ (kJ mol^{-1})
H_2O@Ru-MFI-95	95	2.355556	6.36	8	14.8	7	8	25.5
H_2O@Rh-MFI-95	95	2.355556	6.36	9	14.8	8	8	21.8
H_2O@Pd-MFI-95	95	2.355556	6.36	10	14.8	10	8	16.74
H_2O@Ag-MFI-95	95	2.355556	6.36	11	14.8	10	8	11.3
H_2O@Cd-MFI-95	95	2.355556	6.36	12	14.8	10	8	6.07
H_2O@Ta-MFI-95	95	2.355556	6.36	5	14.8	3	8	36
H_2O@W-MFI-95	95	2.355556	6.36	6	14.8	4	8	35
H_2O@Re-MFI-95	95	2.355556	6.36	7	14.8	5	8	33.05
H_2O@Os-MFI-95	95	2.355556	6.36	8	14.8	6	8	29.29
H_2O@Ir-MFI-95	95	2.355556	6.36	9	14.8	7	8	26.36
H_2O@Pt-MFI-95	95	2.355556	6.36	10	14.8	9	8	19.66
H_2O@Pb-MFI-95	95	2.355556	6.36	14	14.8	0	8	4.77
CO_2@Sc-MFI-47	47	1.927273	6.36	3	26.3	1	16	16.11
CO_2@Ti-MFI-47	47	1.927273	6.36	4	26.3	2	16	18.6
CO_2@V-MFI-47	47	1.927273	6.36	5	26.3	3	16	20.8
CO_2@Cr-MFI-47	47	1.927273	6.36	6	26.3	5	16	20
CO_2@Mn-MFI-47	47	1.927273	6.36	7	26.3	5	16	14.64
CO_2@Fe-MFI-47	47	1.927273	6.36	8	26.3	6	16	13.8
CO_2@Co-MFI-47	47	1.927273	6.36	9	26.3	7	16	16.19
CO_2@Ni-MFI-47	47	1.927273	6.36	10	26.3	8	16	17.2
CO_2@Cu-MFI-47	47	1.927273	6.36	11	26.3	10	16	13.14

(*To be continued*)

System	Si/Al	PLD/d	PLD (Å)	N_G	Polar ($\times 10^{-25} cm^3$)	Nd	Np	$\Delta_{fus}H$ (kJ mol^{-1})
CO_2@Y-MFI-47	47	1.927273	6.36	3	26.3	1	16	17.5
CO_2@Zr-MFI-47	47	1.927273	6.36	4	26.3	2	16	21
CO_2@Nb-MFI-47	47	1.927273	6.36	5	26.3	4	16	26.9
CO_2@Mo-MFI-47	47	1.927273	6.36	6	26.3	5	16	36
CO_2@Ru-MFI-47	47	1.927273	6.36	8	26.3	7	16	25.5
CO_2@Rh-MFI-47	47	1.927273	6.36	9	26.3	8	16	21.8
CO_2@Pd-MFI-47	47	1.927273	6.36	10	26.3	10	16	16.74
CO_2@Ag-MFI-47	47	1.927273	6.36	11	26.3	10	16	11.3
CO_2@Cd-MFI-47	47	1.927273	6.36	12	26.3	10	16	6.07
CO_2@Ta-MFI-47	47	1.927273	6.36	5	26.3	3	16	36
CO_2@W-MFI-47	47	1.927273	6.36	6	26.3	4	16	35
CO_2@Re-MFI-47	47	1.927273	6.36	7	26.3	5	16	33.05
CO_2@Os-MFI-47	47	1.927273	6.36	8	26.3	6	16	29.29
CO_2@Ir-MFI-47	47	1.927273	6.36	9	26.3	7	16	26.36
CO_2@Pt-MFI-47	47	1.927273	6.36	10	26.3	9	16	19.66
CO_2@Pb-MFI-47	47	1.927273	6.36	14	26.3	0	16	4.77
CO_2@Sc-MFI-47	47	1.766667	6.36	3	17.7	1	10	16.11
N_2@Ti-MFI-47	47	1.766667	6.36	4	17.7	2	10	18.6
N_2@V-MFI-47	47	1.766667	6.36	5	17.7	3	10	20.8
N_2@Cr-MFI-47	47	1.766667	6.36	6	17.7	5	10	20
N_2@Mn-MFI-47	47	1.766667	6.36	7	17.7	5	10	14.64

System	d (Å)	R_{DLS}	AN	χ	R (pm)	IE (kJ mol^{-1})	E_b (eV)	I
CO$_2$@Sc-MFI-95	3.3	0.002	21	1.36	144	633.1	-0.76	0.43
CO$_2$@Ti-MFI-95	3.3	0.002	22	1.54	132	658.8	-0.78	0.35
CO$_2$@V-MFI-95	3.3	0.002	23	1.63	122	650.9	-1.28	0.33
CO$_2$@Cr-MFI-95	3.3	0.002	24	1.66	118	652.9	-0.25	0.18
CO$_2$@Mn-MFI-95	3.3	0.002	25	1.55	117	717.3	-1.13	0.29
CO$_2$@Fe-MFI-95	3.3	0.002	26	1.83	117	762.5	-0.76	0.22
CO$_2$@Co-MFI-95	3.3	0.002	27	1.88	116	760.4	-0.47	0.19
CO$_2$@Ni-MFI-95	3.3	0.002	28	1.91	115	737.1	-0.45	0.24
CO$_2$@Cu-MFI-95	3.3	0.002	29	1.9	117	745.5	-0.33	0.20
CO$_2$@Zn-MFI-95	3.3	0.002	30	1.65	125	906.4	-0.38	0.25
CO$_2$@Y-MFI-95	3.3	0.002	39	1.22	162	600	-0.98	0.37
CO$_2$@Zr-MFI-95	3.3	0.002	40	1.33	145	640.1	-0.57	0.39
CO$_2$@Nb-MFI-95	3.3	0.002	41	1.6	134	652.1	-1.41	0.58
CO$_2$@Mo-MFI-95	3.3	0.002	42	2.16	130	684.3	-1.53	0.36
CO$_2$@Ru-MFI-95	3.3	0.002	44	2.2	125	710.2	-0.30	0.20
CO$_2$@Rh-MFI-95	3.3	0.002	45	2.28	125	719.7	-0.08	0.18
CO$_2$@Pd-MFI-95	3.3	0.002	46	2.2	128	804.4	-0.33	0.19
CO$_2$@Ag-MFI-95	3.3	0.002	47	1.93	134	731	-0.56	0.24
CO$_2$@Cd-MFI-95	3.3	0.002	48	1.69	148	867.8	-0.35	0.23
CO$_2$@Ta-MFI-95	3.3	0.002	73	1.5	134	761	-2.09	0.53
CO$_2$@W-MFI-95	3.3	0.002	74	2.36	130	770	-1.63	0.32

(*To be continued*)

System	d (Å)	R_{DLS}	AN	χ	R (pm)	IE (kJ mol^{-1})	E_b (eV)	I
CO_2@Os-MFI-95	3.3	0.002	76	2.2	126	840	-0.69	0.30
CO_2@Ir-MFI-95	3.3	0.002	77	2.2	127	880	-0.32	0.24
CO_2@Pt-MFI-95	3.3	0.002	78	2.28	130	870	-0.32	0.16
CO_2@Pb-MFI-95	3.3	0.002	82	1.8	147	715.6	-0.45	0.25
N_2@Sc-MFI-95	3.6	0.002	21	1.36	144	633.1	0.03	0.19
N_2@Ti-MFI-95	3.6	0.002	22	1.54	132	658.8	-0.94	0.42
N_2@V-MFI-95	3.6	0.002	23	1.63	122	650.9	-1.70	0.50
N_2@Cr-MFI-95	3.6	0.002	24	1.66	118	652.9	-1.41	0.56
N_2@Mn-MFI-95	3.6	0.002	25	1.55	117	717.3	-1.62	0.48
N_2@Fe-MFI-95	3.6	0.002	26	1.83	117	762.5	-1.65	0.54
N_2@Co-MFI-95	3.6	0.002	27	1.88	116	760.4	-1.52	0.54
N_2@Ni-MFI-95	3.6	0.002	28	1.91	115	737.1	-1.15	0.48
N_2@Cu-MFI-95	3.6	0.002	29	1.9	117	745.5	-1.07	0.42
N_2@Zn-MFI-95	3.6	0.002	30	1.65	125	906.4	-0.37	0.25
N_2@Y-MFI-95	3.6	0.002	39	1.22	162	600	0.58	0.08
N_2@Zr-MFI-95	3.6	0.002	40	1.33	145	640.1	-0.60	0.40
N_2@Nb-MFI-95	3.6	0.002	41	1.6	134	652.1	-0.19	0.17
N_2@Mo-MFI-95	3.6	0.002	42	2.16	130	684.3	-1.77	0.46
N_2@Ru-MFI-95	3.6	0.002	44	2.2	125	710.2	-1.31	0.54
N_2@Rh-MFI-95	3.6	0.002	45	2.28	125	719.7	-1.02	0.46
N_2@Pd-MFI-95	3.6	0.002	46	2.2	128	804.4	-0.98	0.36

System	d (Å)	R_{DLS}	AN	χ	R (pm)	IE (kJ mol^{-1})	E_b (eV)	I
N$_2$@Cd-MFI-95	3.6	0.002	48	1.69	148	867.8	-0.49	0.27
N$_2$@Ta-MFI-95	3.6	0.002	73	1.5	134	761	-1.67	0.35
N$_2$@W-MFI-95	3.6	0.002	74	2.36	130	770	-2.25	0.59
N$_2$@Re-MFI-95	3.6	0.002	75	1.9	128	760	-1.65	0.43
N$_2$@Os-MFI-95	3.6	0.002	76	2.2	126	840	-1.30	0.55
N$_2$@Ir-MFI-95	3.6	0.002	77	2.2	127	880	-1.12	0.53
N$_2$@Pt-MFI-95	3.6	0.002	78	2.28	130	870	-1.49	0.50
N$_2$@Pb-MFI-95	3.6	0.002	82	1.8	147	715.6	-0.49	0.26
H$_2$O@Sc-MFI-95	2.7	0.002	21	1.36	144	633.1	-0.63	0.38
H$_2$O@Ti-MFI-95	2.7	0.002	22	1.54	132	658.8	-0.34	0.23
H$_2$O@V-MFI-95	2.7	0.002	23	1.63	122	650.9	-0.66	0.18
H$_2$O@Cr-MFI-95	2.7	0.002	24	1.66	118	652.9	-0.67	0.27
H$_2$O@Mn-MFI-95	2.7	0.002	25	1.55	117	717.3	-0.85	0.22
H$_2$O@Fe-MFI-95	2.7	0.002	26	1.83	117	762.5	-0.81	0.23
H$_2$O@Co-MFI-95	2.7	0.002	27	1.88	116	760.4	-0.85	0.27
H$_2$O@Ni-MFI-95	2.7	0.002	28	1.91	115	737.1	-0.59	0.28
H$_2$O@Cu-MFI-95	2.7	0.002	29	1.9	117	745.5	-0.98	0.38
H$_2$O@Zn-MFI-95	2.7	0.002	30	1.65	125	906.4	-1.09	0.51
H$_2$O@Y-MFI-95	2.7	0.002	39	1.22	162	600	-1.39	0.55
H$_2$O@Zr-MFI-95	2.7	0.002	40	1.33	145	640.1	0.04	0.21
H$_2$O@Nb-MFI-95	2.7	0.002	41	1.6	134	652.1	-0.56	0.25

(*To be continued*)

System	d (Å)	R_{DLS}	AN	χ	R (pm)	IE (kJ mol^{-1})	E_b (eV)	I
H_2O@Ru-MFI-95	2.7	0.002	44	2.2	125	710.2	-0.59	0.26
H_2O@Rh-MFI-95	2.7	0.002	45	2.28	125	719.7	-0.78	0.36
H_2O@Pd-MFI-95	2.7	0.002	46	2.2	128	804.4	-1.19	0.45
H_2O@Ag-MFI-95	2.7	0.002	47	1.93	134	731	-1.17	0.44
H_2O@Cd-MFI-95	2.7	0.002	48	1.69	148	867.8	-1.11	0.50
H_2O@Ta-MFI-95	2.7	0.002	73	1.5	134	761	-0.61	0.12
H_2O@W-MFI-95	2.7	0.002	74	2.36	130	770	-0.38	0.09
H_2O@Re-MFI-95	2.7	0.002	75	1.9	128	760	-0.30	0.11
H_2O@Os-MFI-95	2.7	0.002	76	2.2	126	840	0.03	0.15
H_2O@Ir-MFI-95	2.7	0.002	77	2.2	127	880	-0.29	0.23
H_2O@Pt-MFI-95	2.7	0.002	78	2.28	130	870	-1.11	0.34
H_2O@Pb-MFI-95	2.7	0.002	82	1.8	147	715.6	-1.15	0.50
CO_2@Sc-MFI-47	3.3	0.002	21	1.36	144	633.1	-0.93	0.28
CO_2@Ti-MFI-47	3.3	0.002	22	1.54	132	658.8	-1.06	0.28
CO_2@V-MFI-47	3.3	0.002	23	1.63	122	650.9	-0.78	0.23
CO_2@Cr-MFI-47	3.3	0.002	24	1.66	118	652.9	-0.03	0.16
CO_2@Mn-MFI-47	3.3	0.002	25	1.55	117	717.3	-0.12	0.14
CO_2@Fe-MFI-47	3.3	0.002	26	1.83	117	762.5	-0.36	0.20
CO_2@Co-MFI-47	3.3	0.002	27	1.88	116	760.4	-0.39	0.21
CO_2@Ni-MFI-47	3.3	0.002	28	1.91	115	737.1	-0.29	0.20
CO_2@Cu-MFI-47	3.3	0.002	29	1.9	117	745.5	-0.41	0.22

System	d (Å)	R_{DLS}	AN	χ	R (pm)	IE (kJ mol^{-1})	E_b (eV)	I
CO_2@Y-MFI-47	3.3	0.002	39	1.22	162	600	-1.32	0.36
CO_2@Zr-MFI-47	3.3	0.002	40	1.33	145	640.1	-2.35	0.57
CO_2@Nb-MFI-47	3.3	0.002	41	1.6	134	652.1	-2.15	0.49
CO_2@Mo-MFI-47	3.3	0.002	42	2.16	130	684.3	-1.34	0.30
CO_2@Ru-MFI-47	3.3	0.002	44	2.2	125	710.2	-0.34	0.15
CO_2@Rh-MFI-47	3.3	0.002	45	2.28	125	719.7	-0.31	0.22
CO_2@Pd-MFI-47	3.3	0.002	46	2.2	128	804.4	-0.16	0.25
CO_2@Ag-MFI-47	3.3	0.002	47	1.93	134	731	-0.48	0.27
CO_2@Cd-MFI-47	3.3	0.002	48	1.69	148	867.8	-0.78	0.20
CO_2@Ta-MFI-47	3.3	0.002	73	1.5	134	761	-0.51	0.17
CO_2@W-MFI-47	3.3	0.002	74	2.36	130	770	-0.36	0.16
CO_2@Re-MFI-47	3.3	0.002	75	1.9	128	760	-1.14	0.26
CO_2@Os-MFI-47	3.3	0.002	76	2.2	126	840	-0.31	0.15
CO_2@Ir-MFI-47	3.3	0.002	77	2.2	127	880	-0.22	0.21
CO_2@Pt-MFI-47	3.3	0.002	78	2.28	130	870	-0.14	0.29
CO_2@Pb-MFI-47	3.3	0.002	82	1.8	147	715.6	-0.32	0.27
CO_2@Sc-MFI-47	3.6	0.002	21	1.36	144	633.1	-1.04	0.31
N_2@Ti-MFI-47	3.6	0.002	22	1.54	132	658.8	-1.22	0.33
N_2@V-MFI-47	3.6	0.002	23	1.63	122	650.9	-1.32	0.39
N_2@Cr-MFI-47	3.6	0.002	24	1.66	118	652.9	-0.86	0.36
N_2@Mn-MFI-47	3.6	0.002	25	1.55	117	717.3	-1.01	0.34
N_2@Fe-MFI-47	3.6	0.002	26	1.83	117	762.5	-0.70	0.28

Supplementary Table 4. Data records in zeolite dataset

| Feature	Feature distribution	Data sources
Silicon aluminum ratios (Si/Al)		Model construction
Pore limiting diameter (PLD)		Zeolite structure databases [1]
d-shell valence electron numbers (Nd)		Periodic table of elements
The number of the group in the periodic table (N_G)		Periodic table of elements
The enthalpy of fusion ($\Delta_{fus}H$)		CRC handbook of chemistry and physics [12]
Polarizability (Polar)		CRC handbook of chemistry and physics [12]
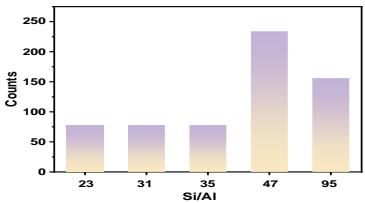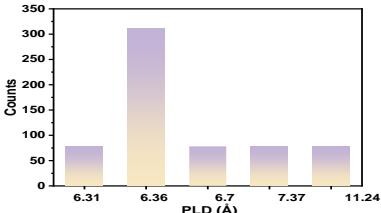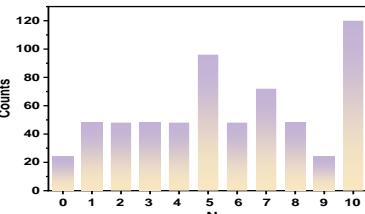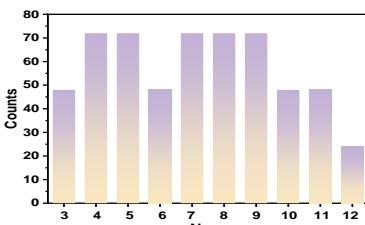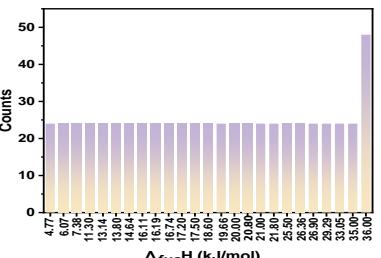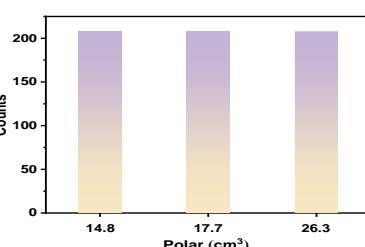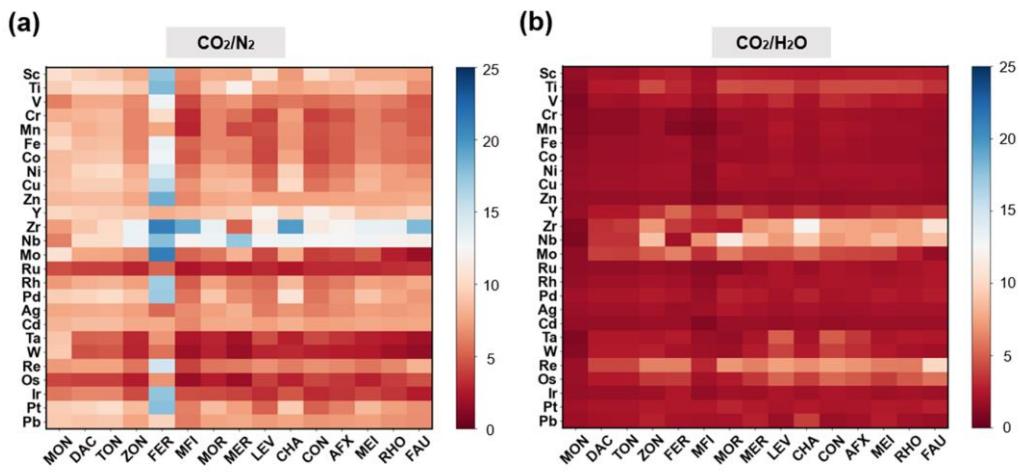
Supplementary Figure 9. The adsorption selectivity heat map of (a) CO_2/N_2, (b) CO_2/H_2O adsorption separation.

Supplementary Table 5. Parameters of the XGBoost algorithm for predicting the formation energy.

Parameter	value
learning_rate	0.11
max_dapth	16
n_estimators	580
min_child_weight	5
subsample	0.84
colsample_bytree	0.8
reg_alpha	0.8
reg_lambda	0.3

Supplementary Table 6. Parameters of the ExtraTree algorithm for predicting the binding energy and adsorption strength index.

Parameter	value
min_samples_leaf	1
max_dapth	12
n_estimators	320
random_state	7
min_samples_split	4
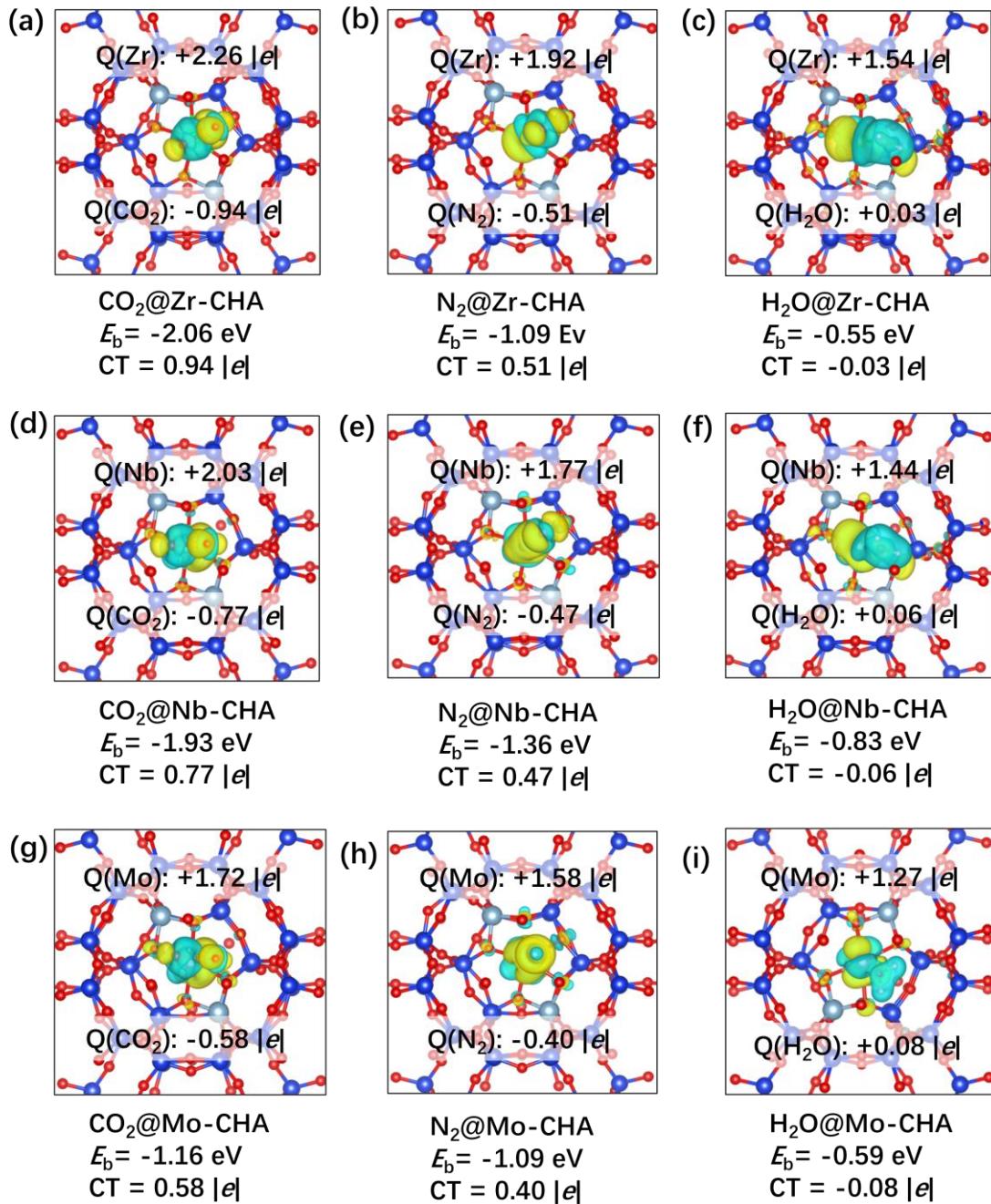
Supplementary Figure 10. The charge density difference of the (a) CO_2, (b) N_2, (c) H_2O on Zr-CHA-47; (d) CO_2, (e) N_2, (f) H_2O on Nb-CHA-47; (g) CO_2, (h) N_2, (i) H_2O on Mo-CHA-47.

REFERENCES

1. Baerlocher, C.; McCusker, L. Database of Zeolite Structures. Available from: http://www.iza structure.org/databases/. [Last accessed on 27 July 2023]

2. D'Angelo GM, Rao D, Gu CC. Combining least absolute shrinkage and selection operator (LASSO) and principal-components analysis for detection of gene-gene interactions in genome-wide association studies. *BMC Proc* 2009**;**Suppl 7:S62. [DOI:10.1186/1753-6561-3-s7-s62]

3. SMOLA AJ, SCHOLKOPF B. A tutorial on support vector regression. *Stat and Comput* 2004:199–222. [DOI:10.1023/B:STCO.0000035301.49549.88]

4. Rodríguez JJ, Quintana G, Bustillo A, Ciurana J. A decision-making tool based on decision trees for roughness prediction in face milling. *Int J Comput Integ M* 2016**;**9:943-957. [DOI:10.1080/0951192x.2016.1247991]

5. Patrick EA FF. A generalized k-nearest neighbor rule. *Inf Control* 1970:128. [DOI:DOI: 10.1016/S0019-9958(70)90081-1]

6. Hoerl RW. Ridge Regression: A Historical Context. *Technometrics* 2020**;**4:420-425. [DOI:10.1080/00401706.2020.1742207]

7. Schumann TEW. XXVI. The principles of a mechanical method for calculating regression equations and multiple correlation coefficients and for the solution of simultaneous linear equations. *Lond Edinb Dublin Philos Mag J Sci* 1940**;**194:258-273. [DOI:10.1080/14786444008521266]

8. Abiodun OI, Kiru MU, Jantan A, et al. Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition. *IEEE Access* 2019:158820-158846. [DOI:10.1109/access.2019.2945545]

9. Guo Y, Du L, Wei D, Li C. Robust SAR Automatic Target Recognition Via Adversarial Learning. *IEEE J Sel Top Appl Earth Obs Remote Sens* 2021;716-729. [DOI:10.1109/jstars.2020.3039235]

10. Priyadarshi R, Panigrahi A, Routroy S, Garg GK. Demand forecasting at retail stage for selected vegetables: a performance analysis. *J Model Manag* 2019**;**4:1042-1063. [DOI:10.1108/jm2-11-2018-0192]

11. van der Walt S, Schonberger JL, Nunez-Iglesias J, et al. scikit-image: image processing in Python. *PeerJ* 2014:e453. [DOI:10.7717/peerj.453]

12. Tomaszewski R. Citations to chemical resources in scholarly articles: CRC Handbook of Chemistry and Physics and The Merck Index. *Scientometrics* 2017;3:1865-1879. [DOI: 10.1007/s11192-017-2437-4]