Supplementary Information (SI) for Journal of Materials Chemistry A.

Supporting Information

Conformal Active Learning Aided Screening of Ligand Protected Cu-Nanoclusters for CO_2 Reduction Reactions

Diptendu Roy,[†] Amitabha Das,[†] Biswarup Pathak,[†,*]

[†]Department of Chemistry, Indian Institute of Technology Indore, Indore 453552, India

*Email: biswarup@iiti.ac.in

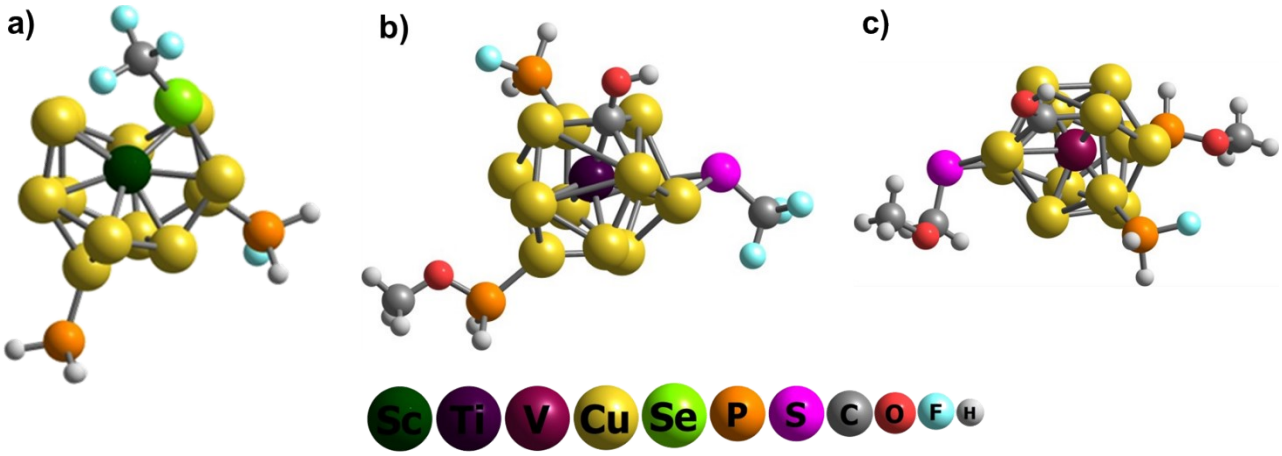

Figure S1. Optimized structures of a) NC with Sc core, b) *COH adsorbed NC with Ti core, and c) *COH adsorbed NC with V core.

Text S1: Calculation for all possible combinations of nanocluster catalysts

We have considered 9 core atom dopants and 9 ligands.

All possible combinations with 3 different ligands = C_3^9 = 84, and possible adsorption sites in between two ligands = 3.

Combinations with two same ligands and one different ligand = 2 × C_2^9 = 2 × 36 = 72, and possible adsorption sites in between two ligands = 2.

Combinations with three same ligands = 9, and possible adsorption sites in between two ligands = 1.

Total possible combinations = 84 × 9 × 3 + 72 × 9 × 2 + 9 × 9 × 1 = 3645

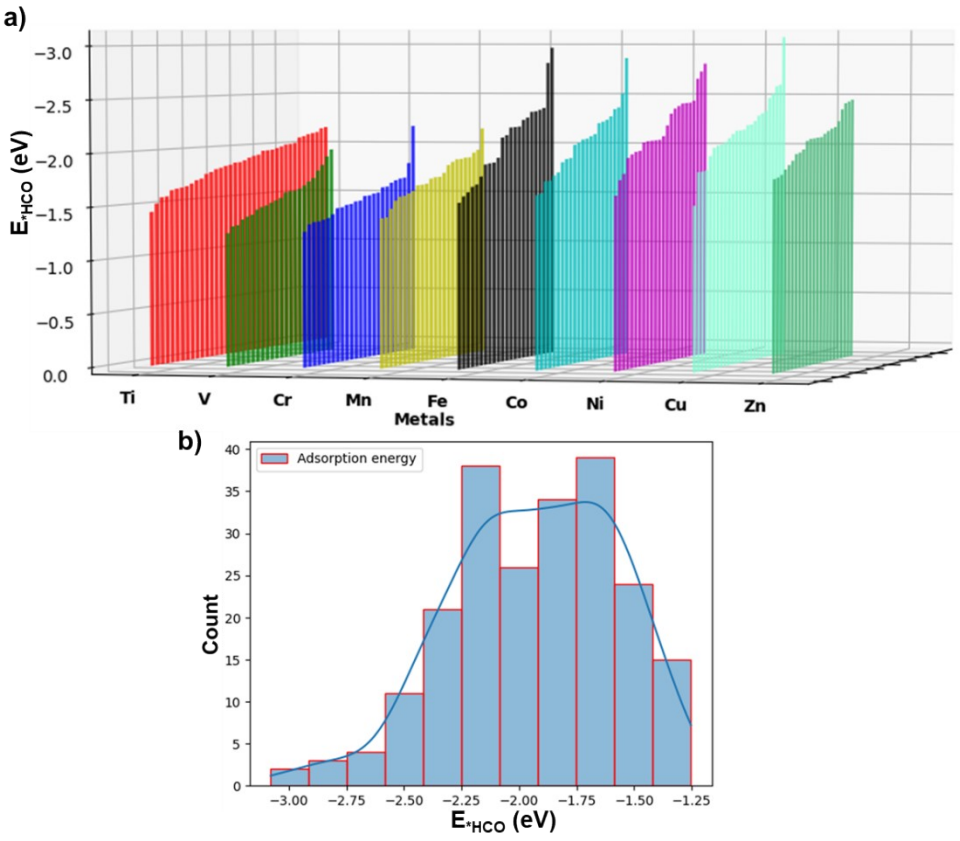

Figure S2. a) The 3D distribution plot of E_{*HCO} with respect to core dopant metals and b) the frequency distribution of E_{*HCO}.

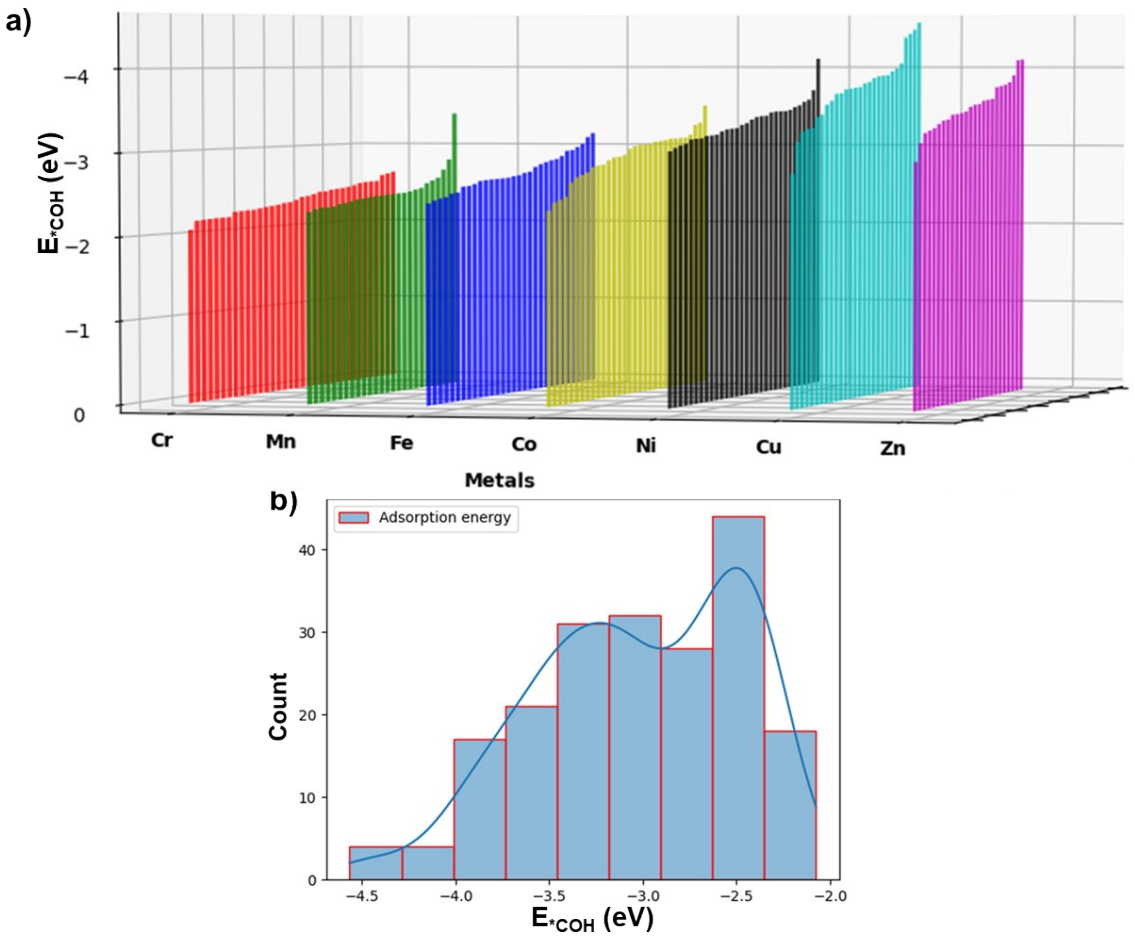

Figure S3. a) The 3D distribution plot of E_{*COH} with respect to core dopant metals and b) the frequency distribution of E_{*COH}.

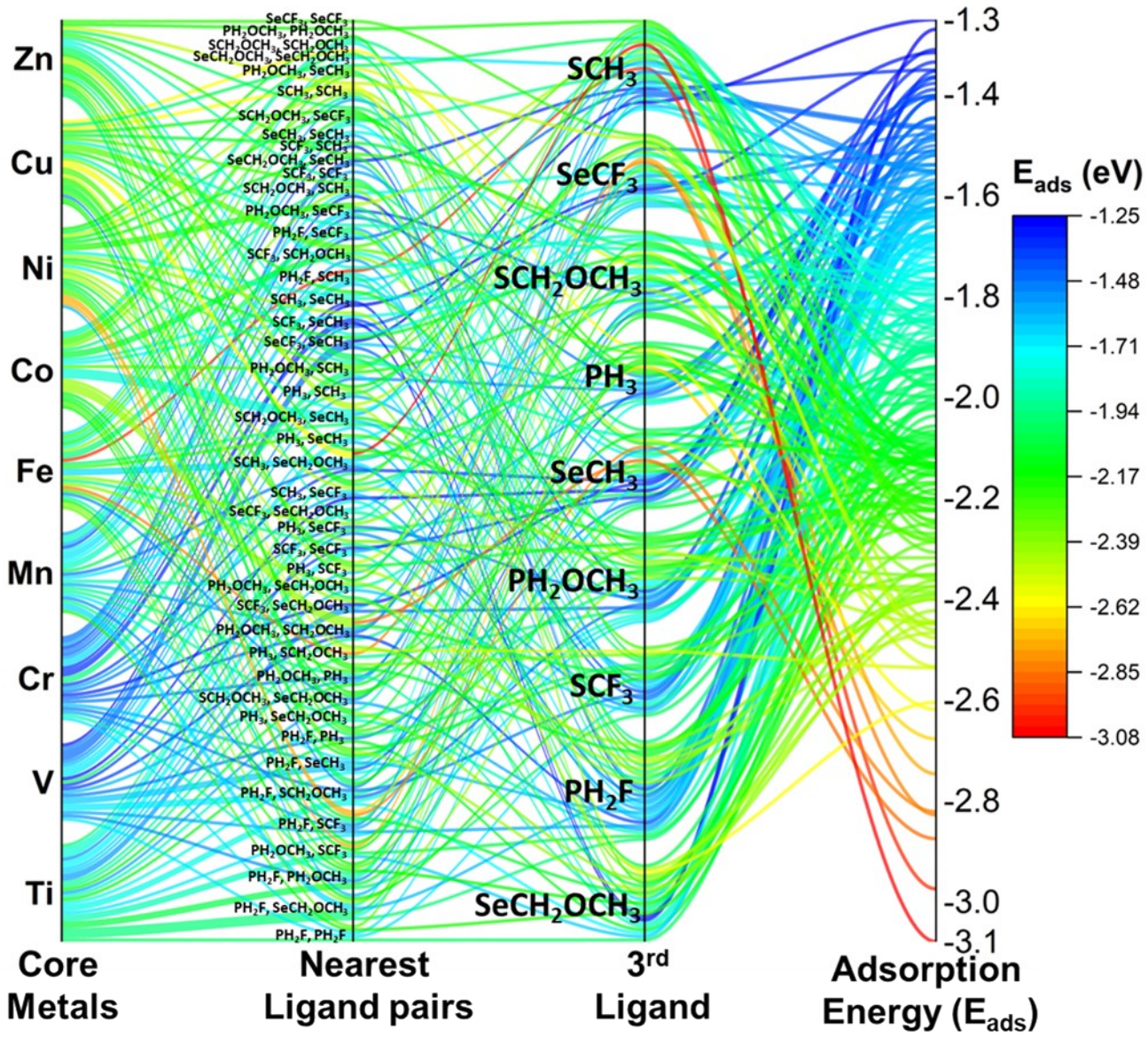

Figure S4. A parallel plot showing the E_{*HCO} with respect to the core metal, ligand combinations and their positions.

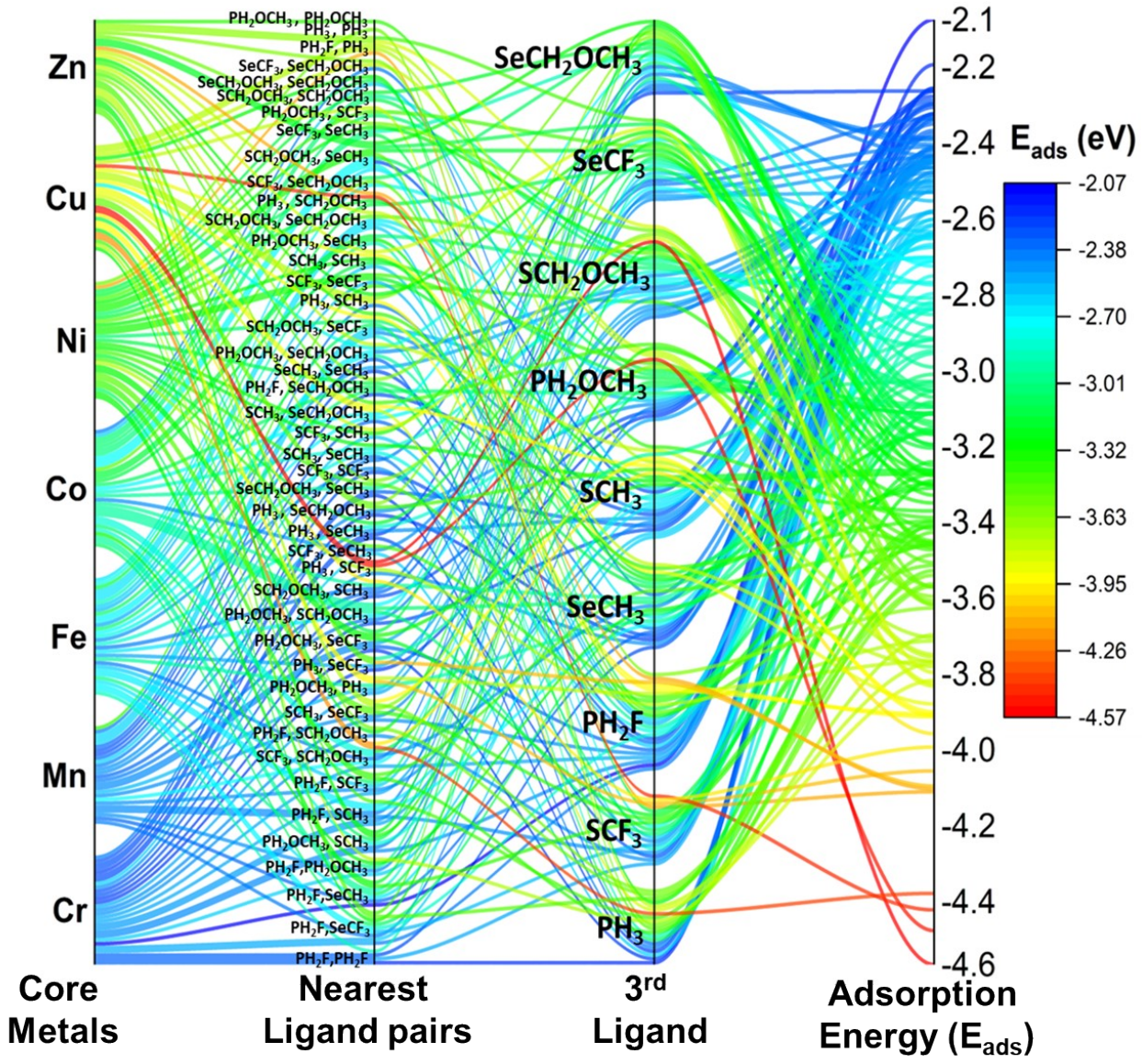

Figure S5. A parallel plot showing the E_{*COH} with respect to the core metal, ligand combinations and their positions.

Table S1: Detail description of all the features considered in the study.

Features	Description
Vol1, Vol2, Vol3	Volume of three protective ligands in nanocluster
Chare1, Charge2, Charge3	Charge of the substituents of three protective ligands in nanocluster
Core_d_band	d-band center of core doped bare nanocluster
Feature1 (For 1st ligand), Feature2 (For 2nd ligand)	(Charge of the substituents + Unsaturation in central atom of the ligand)/ (P orbital radii of ligands central atom)
lig1_1st_en, lig2_1st_en	Electronegativity of central atom of two ligands nearest to the adsorption site
lig1_others_en, lig2_others_en	Avearge electronegativity of the atoms other than central atom involved in two ligands nearest to the adsorption site
lig1_1st_dp, lig2_1st_dp	Dipole polarizability of central atom of two ligands nearest to the adsorption site
lig1_others_dp, lig2_others_dp	Avearge dipole polarizability of the atoms other than central atom involved in two ligands nearest to the adsorption site
1st_P_radii, 2nd_P_radii	P orbital radii of the central atom of two ligands nearest to the adsorption site
1st_p_val, 2nd_p_val	Valence P electron of the central atom of two ligands nearest to the adsorption site
1st_unsaturation, 2nd_unsaturation	Unsaturation in the central atom of two ligands nearest to the adsorption site
metal_core_EN_diff	Electronegativity difference between core and shell metal (Cu)
metal_core_dp_diff	Dipole polarizability difference between core and shell metal (Cu)
metal_core_d_e_diff	Valence d electron difference between core and shell metal (Cu)
metal_core_s_e_diff	Valence s electron difference between core and shell metal (Cu)
metal_core_IP_diff	Ionization potential difference between core and shell metal (Cu)
metal_core_radii_diff	Atomic radii difference between core and shell metal (Cu)
lig1_core_EN_diff, lig2_core_EN_diff, lig3_core_EN_diff	Average electronegativity difference between liagnd atoms and core metal for all three ligands
lig1_core_dp_diff, lig2_core_dp_diff, lig3_core_dp_diff	Average dipole polarizability difference between liagnd atoms and core metal for all three ligands

Features	Description
lig1_core_val_diff, lig2_core_val_diff, lig3_core_val_diff	Average valence electron difference between ligand atoms and core metal for all three ligands
lig1_core_IP_diff, lig2_core_IP_diff, lig3_core_IP_diff	Average ionization potential difference between ligand atoms and core metal for all three ligands
lig1_metal_EN_diff, lig2_metal_EN_diff, lig3_metal_EN_diff	Average electronegativity difference between ligand atoms and shell metal (Cu) for all three ligands
lig1_metal_dp_diff, lig2_metal_dp_diff, lig3_metal_dp_diff	Average dipole polarizability difference between ligand atoms and shell metal (Cu) for all three ligands
lig1_metal_val_diff, lig2_metal_val_diff, lig3_metal_val_diff	Average valence electron difference between ligand atoms and shell metal for all three ligands
lig1_metal_IP_diff, lig2_metal_IP_diff, lig3_metal_IP_diff	Average ionization potential difference between ligand atoms and shell metal for all three ligands

Table S2. Best-fitted hyperparameters for all the target descriptors with their corresponding train and cross-validated test RMSE.

Target Descriptor	Best-fitted ML model with optimized hyperparameters	Train RMSE (eV)	Cross-validation RMSE (eV)
E_{*CO}	**RFR** (n_estimators=100, min_samples_split=5, min_samples_leaf=1, max_features='sqrt', max_depth=10, bootstrap=True, random_state=42)	0.11	0.21
E_{*HCO}	**GBR** (n_estimators=100, min_samples_split=15, min_samples_leaf=5, max_depth=15, learning_rate=0.05)	0.05	0.30
E_{*COH}	**RFR** (n_estimators=60, min_samples_split=4, min_samples_leaf=2, max_features='sqrt', max_depth=40, bootstrap=False, random_state=30)	0.08	0.26
$\Delta E_{*CO \rightarrow *COH}$	**GBR** (n_estimators=100, min_samples_split=5, min_samples_leaf=15, max_depth=10, learning_rate=0.1)	0.09	0.27
$\Delta E_{*CO \rightarrow *HCO}$	**GBR** (n_estimators=100, min_samples_split=10, min_samples_leaf=10, max_depth=5, learning_rate=0.05)	0.12	0.26

Text S2. **Conformal prediction and active learning**

Conformal prediction: Conformal prediction comes into the picture due to the failure of various uncertainty quantification methods to provide coverage guarantee of the prediction sets.[1] It can be set of classes in case of classification problems and prediction interval in case of regression problems. The conformal prediction could provide mathematically rigorous guarantees and it is statistically valid. Conformal prediction can be of two types. One is classical conformal prediction i.e., transductive or full conformal prediction which uses the entire dataset, including the new data point, for creating prediction regions. Transductive conformal prediction doesn't split the data and instead refits the model multiple times to produce a prediction region. To get the prediction set for a new data point, the model must be retrained for every new data point.

Another type is inductive conformal prediction in which data is split into train and calibration data and all the scores rely on the calibration data. After splitting the data, it is required to choose a suitable ML model to use as underlying model. That can be chosen using usual process of hyperparameter tuning and evaluate on a test set based on error metrics such as mean absolute error. Then the nonconformity measure is required to define which indicates how rare an observation is compared to rest of the data. With the nonconformity measure we calculate the nonconformity scores on the calibration data. Afterwards, a desired coverage or confidence level has to be chosen and after all this set up, the prediction of unknown dataset can provide various prediction intervals for different unknown data points.

In this study, we have used inductive conformal prediction. To implement the conformal prediction, we have used nonconformist package here. It can use the underlying ML model from sklearn package. It can create nonconformity function using NcFactory of nonconformist package which can only specifically use the algorithms imported from sklearn.

In this study, as we have dealt with regression problems, we have chosen absolute error as nonconformity function. Here, nonconformity function $\alpha = |y_i - \hat{y}_i|$ where, y_i is ith the data point in calibration data and $\hat{y}_i$ is the corresponding ith predicted data point.

Moreover, in nonconformist package, nonconformity functions leverage an additional underlying model that attempts to predict the difficulty of predicting the output of a given test pattern. In this study, we have KNeighborsRegressor as the additional underlying model for normalizing the nonconformity function. Finally, with all this set up, while we have predicted for unknown data sets, it provides us with different prediction intervals for different unknown data points. Here, to identify the most uncertain data, we have used the length of the prediction intervals as the query strategy which is done usually in active learning.

Table S3: Iteratively added data points (catalysts) for E_{*CO}.

Iterations	Core	Ligand pair	3rd ligand	E_{*CO} (eV)
Iteration1	Cr	('$SeCF_3$', '$SeCF_3$')	'SCH_2OCH_3'	-1.12
Iteration 1	Cr	('$SeCF_3$', '$SeCF_3$')	'$SeCH_2OCH_3$'	-0.68
Iteration 2	Cr	('SCF_3','$SeCH_2OCH_3$')	'SCF_3'	-1.04
Iteration 2	Cr	('PH_2OCH_3', 'SCH_2OCH_3')	'$SeCF_3$'	-1.37
Iteration 3	Cr	('PH_2OCH_3', 'SCH_2OCH_3')	'$SeCH_2OCH_3$'	-1.31
Iteration 3	Cr	('PH_2OCH_3', 'SCH_2OCH_3')	PH_2OCH_3'	-1.21
Iteration 4	Cr	('PH_2OCH_3', '$SeCH_2OCH_3$')	PH_2OCH_3'	-1.20
Iteration 4	Cr	('SCF_3', 'SCH_2OCH_3')	'SCF_3'	-1.08
Iteration 5	Mn	('SCH_2OCH_3', 'SCH_3')	'SCH_3'	-1.13
Iteration 5	Mn	('SCH_3', '$SeCH_2OCH_3$')	'SCH_3'	-1.41

Table S4: Train and test MAE, prediction interval coverage probability and mean prediction interval (MPI) for each iteration of E_{*CO}.

Iterations	Train MAE (eV)	Test Avg. CV MAE (eV)	PICP (%)	MPI (eV)
1	0.08	0.17	82	0.59
2	0.06	0.17	82	0.55
3	0.08	0.16	82	0.52
4	0.08	0.16	82	0.44
5	0.07	0.16	82	0.50

Table S5: Iteratively added data points (catalysts) for E_{*HCO}.

Iterations	Core	Ligand pair	3rd ligand	E_{*HCO} (eV)
Iteration1	Ti	('PH_2F', 'SCH_2OCH_3')	'PH_2OCH_3'	-1.81
Iteration 1	Ti	('PH_2F', 'PH_2OCH_3')	'$SeCH_3$'	-2.08
Iteration 1	Ti	('PH_2F', '$SeCH_3$')	'PH_2OCH_3'	-1.77
Iteration 2	Cu	('PH_2F', 'PH_2OCH_3')	'SCH_3'	-2.16
Iteration 2	Ti	('PH_2F', 'PH_3')	'SCF_3'	-2.15
Iteration 2	Co	('PH_2OCH_3', 'PH_2OCH_3')	'PH_2F'	-1.96
Iteration 3	Ti	('PH_2F', 'SCF_3')	'PH_3'	-1.91
Iteration 3	Ti	('PH_3', 'SCH_2OCH_3')	'PH_2F'	-2.01
Iteration 4	Ti	('PH_2F', 'PH_2OCH_3')	'$SeCF_3$'	-2.07
Iteration 4	Ti	('PH_3', 'SCF_3')	'PH_2F'	-1.84
Iteration 5	Ti	('PH_2F', 'SCH_3')	'PH_2OCH_3'	-1.97
Iteration 5	Ti	('PH_3', 'SCH_3')	'PH_2F'	-1.86

Table S6: Train and test MAE, prediction interval coverage probability and MPI for each iteration of E_{*HCO}.

Iterations	Train MAE (eV)	Test Avg. CV MAE (eV)	PICP (%)	MPI (eV)
1	0.04	0.24	82	0.87
2	0.05	0.21	82	0.82
3	0.05	0.20	82	0.72
4	0.05	0.22	82	0.70
5	0.05	0.21	82	0.68

Table S7: Iteratively added data points (catalysts) for E_{*COH}.

Iterations	Core	Ligand pair	3rd ligand	E_{*COH} (eV)
Iteration1	Cr	('PH_2F', 'PH_2F')	'PH_3'	-2.19
Iteration 1	Cr	('PH_2F', 'PH_2F')	'SCF_3'	-2.49
Iteration 2	Cr	('PH_2F', '$SeCF_3$')	'PH_2F'	-2.41
Iteration 2	Cr	('PH_2F', 'PH_3')	'PH_2F'	-2.19
Iteration 3	Cr	('PH_2OCH_3', 'PH_2OCH_3')	'PH_2F'	-2.19
Iteration 3	Cu	('PH_2F', 'SCH_3')	'$SeCH_2OCH_3$'	-3.09
Iteration 4	Cr	('PH_2F', 'PH_2OCH_3')	'PH_3'	-2.18
Iteration 4	Zn	('PH_3', '$SeCF_3$')	'PH_2OCH_3'	-3.59
Iteration 5	Cr	('PH_2F', 'SCF_3')	'PH_2F'	-2.46
Iteration 5	Cr	('PH_2F', 'PH_3')	'PH_2OCH_3'	-2.20

Table S8: Train and test MAE, prediction interval coverage probability and MPI for each iteration of E_{*COH}.

Iterations	Train MAE (eV)	Test Avg. CV MAE (eV)	PICP (%)	MPI (eV)
1	0.07	0.21	82	0.76
2	0.07	0.18	83	0.71
3	0.08	0.18	83	0.63
4	0.07	0.17	83	0.59
5	0.05	0.16	83	0.50

Table S9: Iteratively added data points (catalysts) for $\Delta E_{*CO \rightarrow *HCO}$.

Iterations	Core	Ligand pair	3rd ligand	$\Delta E_{*CO \rightarrow *HCO}$ (eV)
Iteration 1	Ti	('PH_2F', 'SCH_3')	'SCH_2OCH_3'	0.69
Iteration 1	Ti	('PH_2F', '$SeCH_3$')	'SCH_2OCH_3'	0.70
Iteration 2	Ti	('PH_2F', '$SeCH_3$')	'$SeCH_2OCH_3$'	0.69
Iteration 2	Fe	('SCF_3', 'SCF_3')	'SCH_2OCH_3'	1.46
Iteration 3	Fe	('SCF_3', 'SCH_2OCH_3')	'$SeCF_3$'	0.80
Iteration 3	Fe	('SCH_2OCH_3', '$SeCF_3$')	'$SeCF_3$'	0.65
Iteration 4	Cu	('PH_2OCH_3', '$SeCH_2OCH_3$')	'$SeCH_3$'	0.45
Iteration 4	Ni	('PH_3', 'SCH_3')	'SCH_3'	0.47
Iteration 5	Fe	('$SeCH_2OCH_3$', '$SeCH_3$')	'$SeCH_3$'	0.77
Iteration 5	Ni	('PH_2OCH_3', '$SeCH_2OCH_3$')	'$SeCH_3$'	0.42
Iteration 6	Ni	('PH_2OCH_3', 'SCH_2OCH_3')	'SCH_3'	-0.19
Iteration 6	Cu	('PH_2OCH_3', 'SCH_2OCH_3')	'SCH_3'	-0.27

Table S10: Train and test MAE, prediction interval coverage probability and MPI for each iteration of $\Delta E_{*CO \rightarrow *HCO}$.

Iterations	Train MAE (eV)	Test Avg. CV MAE (eV)	PICP (%)	MPI (eV)
1	0.10	0.21	83	0.83
2	0.09	0.20	83	0.78
3	0.06	0.20	83	0.76
4	0.09	0.22	83	0.71
5	0.06	0.21	83	0.68
6	0.07	0.19	83	0.66

Table S11: Iteratively added data points (catalysts) for $\Delta E_{*CO \rightarrow *COH}$.

Iterations	Core	Ligand pair	3rd ligand	$\Delta E_{*CO \rightarrow *COH}$ (eV)
Iteration 1	Cr	('$SeCH_2OCH_3$', '$SeCH_3$')	'$SeCH_3$'	2.06
Iteration 1	Cr	('PH_2F', 'PH_2OCH_3')	'$SeCF_3$'	2.13
Iteration 2	Zn	('PH_2F', '$SeCH_2OCH_3$')	'PH_3'	1.06
Iteration 2	Cr	('PH_2F', 'PH_2F')	'PH_2OCH_3'	2.42
Iteration 3	Cr	('PH_2F', 'PH_2OCH_3')	'SCF_3'	2.13
Iteration 3	Cr	('PH_2F', 'PH_2OCH_3')	'SCH_2OCH_3'	2.35
Iteration 4	Cr	('PH_2OCH_3', '$SeCH_2OCH_3$')	'PH_2OCH_3'	2.26
Iteration 4	Cr	('SCH_3', '$SeCF_3$')	'$SeCF_3$'	1.81
Iteration 5	Cr	('PH_2OCH_3', '$SeCH_2OCH_3$')	'$SeCH_3$'	1.73
Iteration 5	Cr	('SCH_2OCH_3', 'SCH_2OCH_3')	'$SeCH_2OCH_3$'	2.23
Iteration 6	Cr	('SCH_2OCH_3', '$SeCH_2OCH_3$')	'$SeCH_2OCH_3$'	1.67
Iteration 6	Mn	('PH_2OCH_3', 'PH_2OCH_3')	'PH_2F'	1.88

Table S12: Train and test MAE, prediction interval coverage probability and MPI for each iteration of $\Delta E_{*CO \rightarrow *COH}$.

Iterations	Train MAE (eV)	Test Avg. CV MAE (eV)	PICP (%)	MPI (eV)
1	0.08	0.21	84	0.79
2	0.08	0.21	84	0.78
3	0.09	0.20	84	0.67
4	0.06	0.20	84	0.67
5	0.06	0.21	84	0.73
6	0.06	0.21	84	0.73

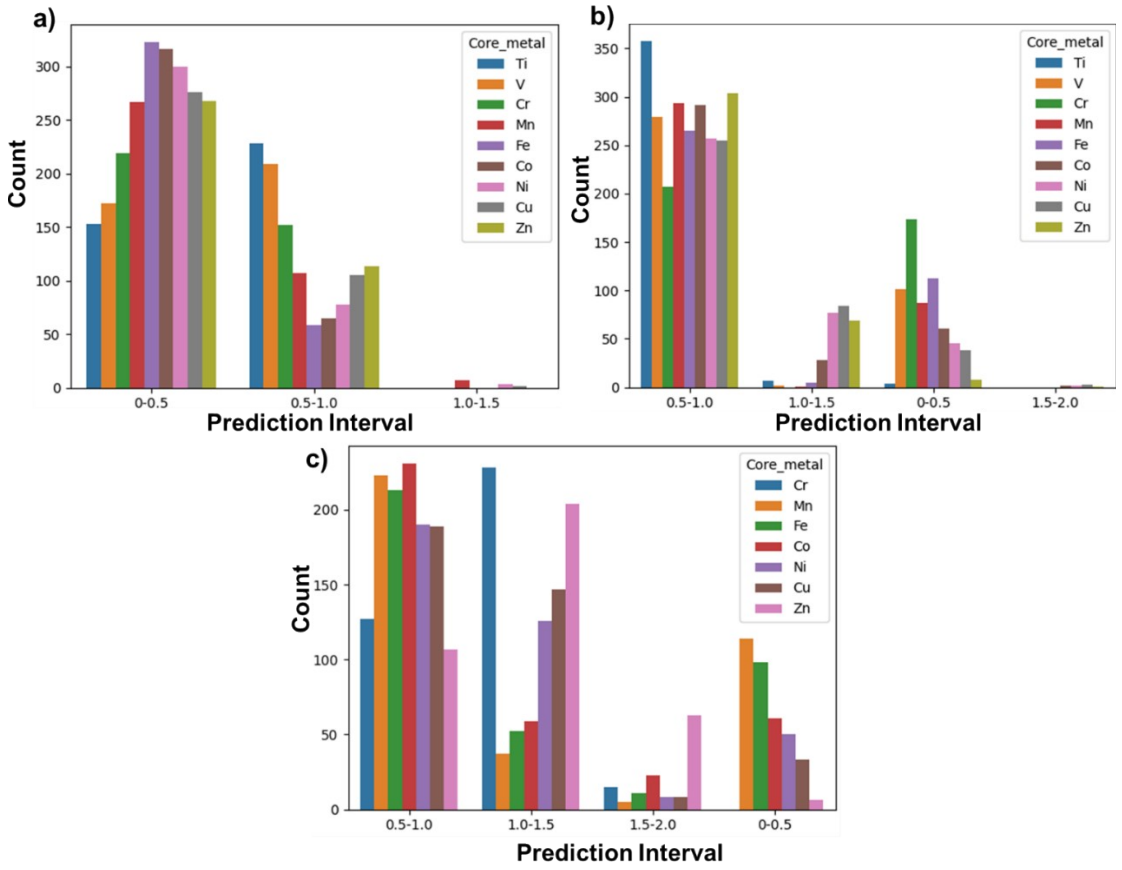

Figure S6. Prediction interval for predicted unknown datasets of a) E_{*CO}, b) E_{*HCO} and c) E_{*COH}.

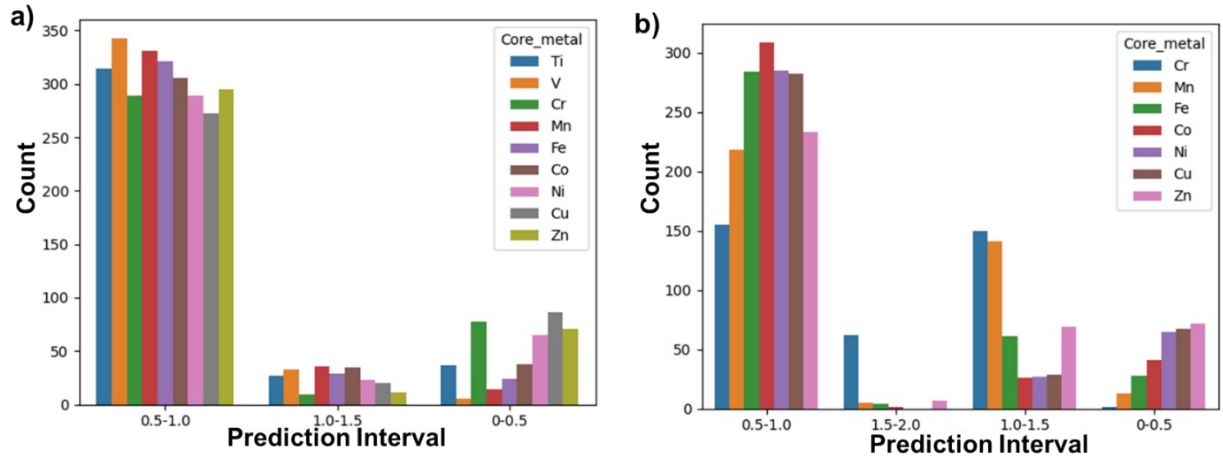

Figure S7. Prediction interval for predicted unknown datasets of a) $\Delta E_{*CO \rightarrow *HCO}$, and b) $\Delta E_{*CO \rightarrow *COH}$.

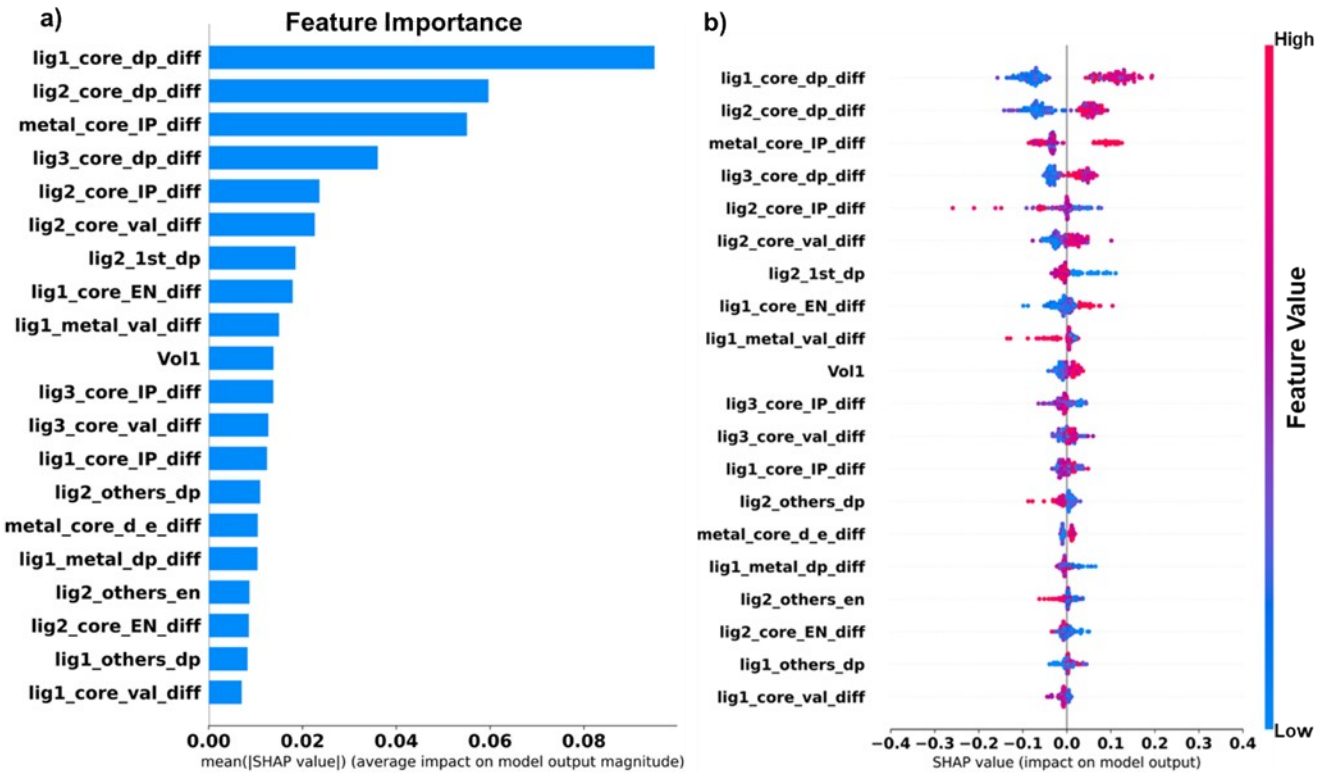

Figure S8. For E_{*HCO} prediction, SHAP plot demonstrates the a) feature importance of higher contributing features among others, and b) bee-swarm plot showing the distribution of feature values with their corresponding SHAP value for most contributing features.

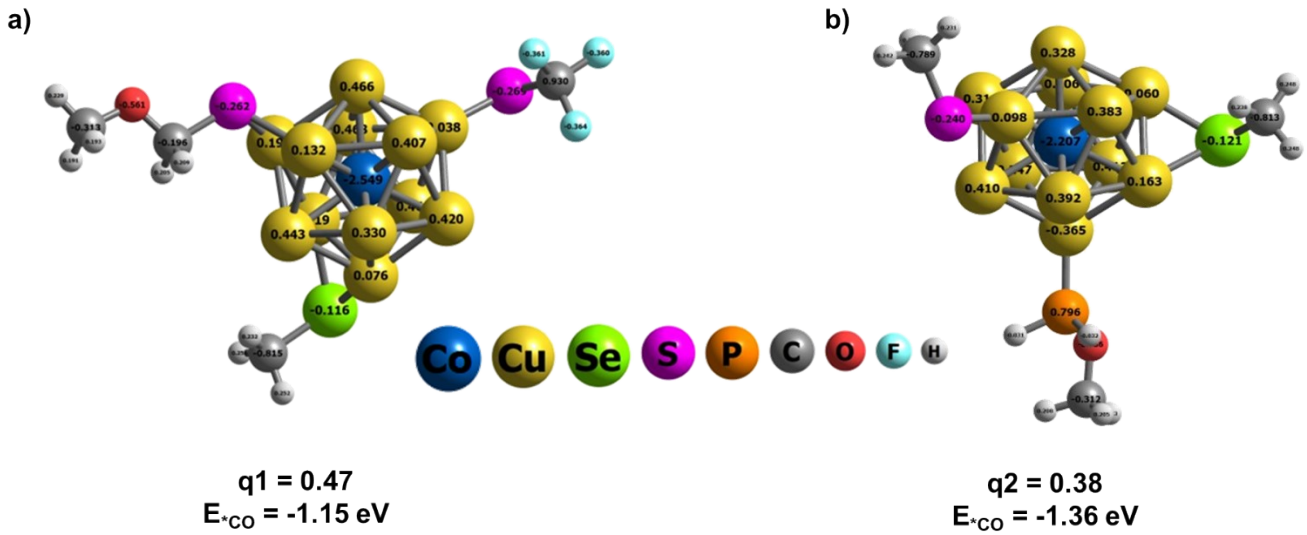

Figure S9. Comparison of NBO charge and *CO adsorption energy for two representative catalysts with same core metal and different ligands. Here, q1 indicates charge of the *CO adsorbed metal site on NC catalysts.

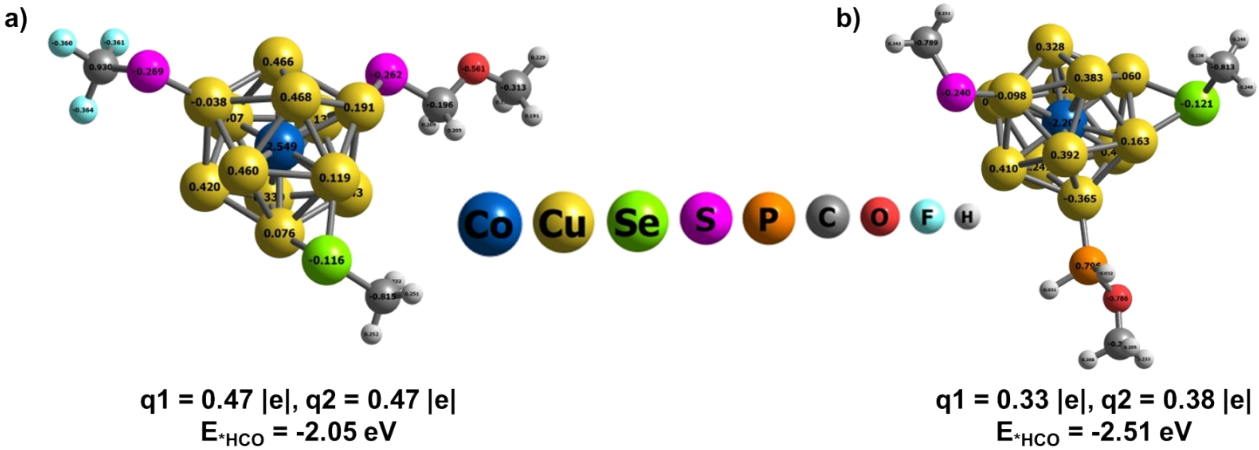

Figure S10. Comparison of NBO charge and *HCO adsorption energy for two representative catalysts with same core metal and different ligands. Here, q1 and q2 indicate charges of the two shell metals belonging to *HCO adsorption site on NC catalysts.

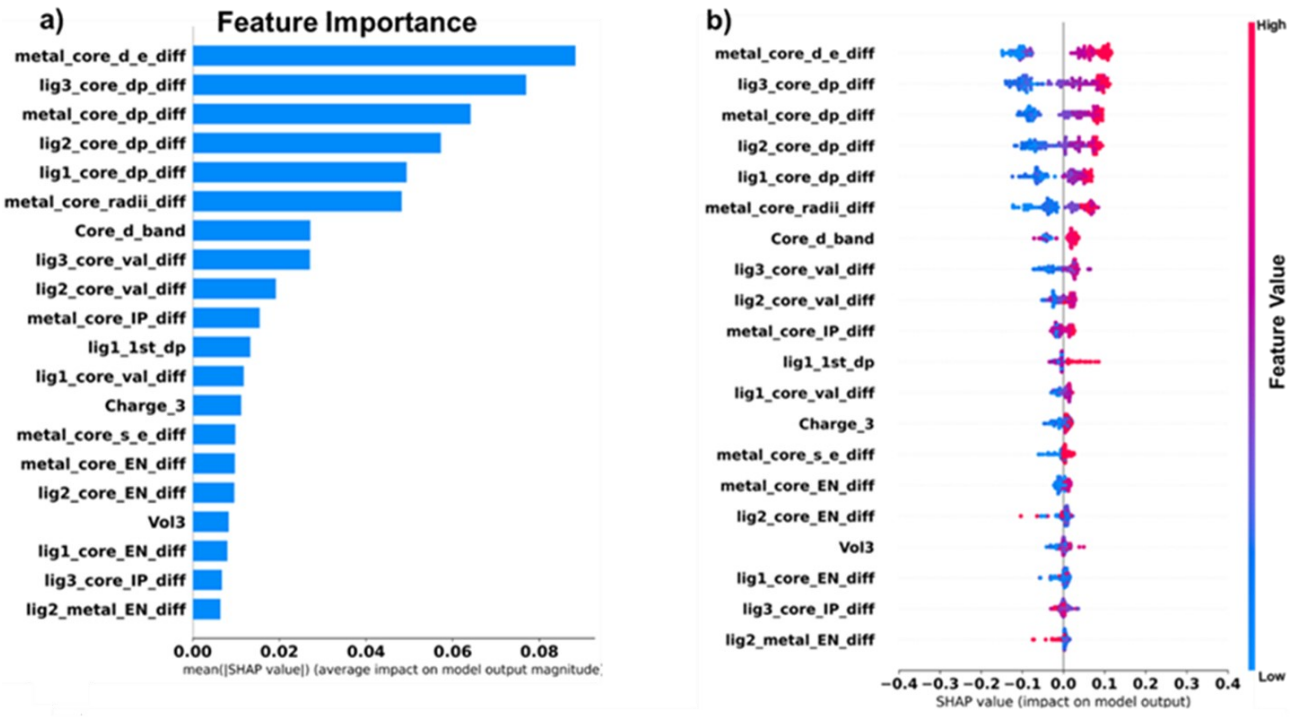

Figure S11. For E_{*COH} prediction, SHAP plot demonstrates the a) feature importance of higher contributing features among others, and b) bee-swarm plot showing the distribution of feature values with their corresponding SHAP value for most contributing features.

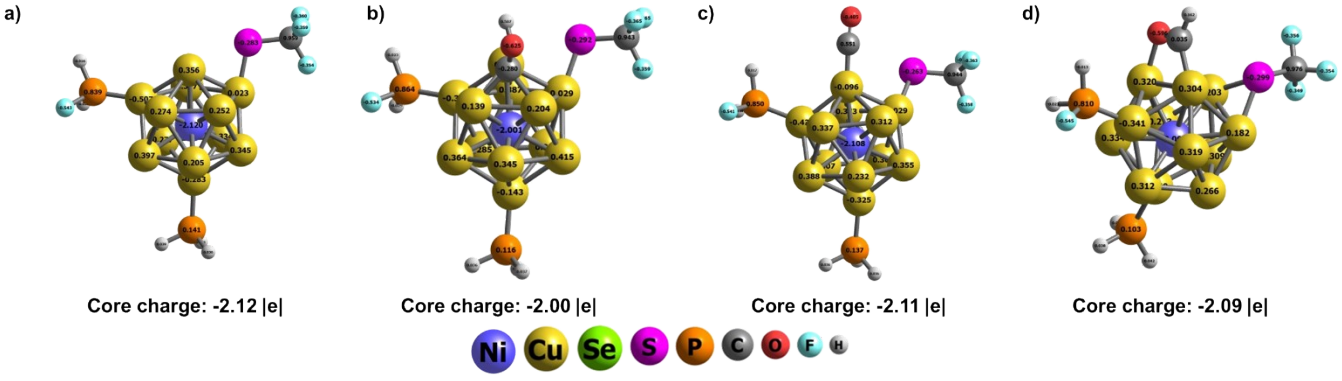

Figure S12. Comparison of core charge of a) a Ni-based representative catalyst without adsorbate with its corresponding b) *COH adsorbed, c) *CO adsorbed, and d) *HCO adsorbed structures.

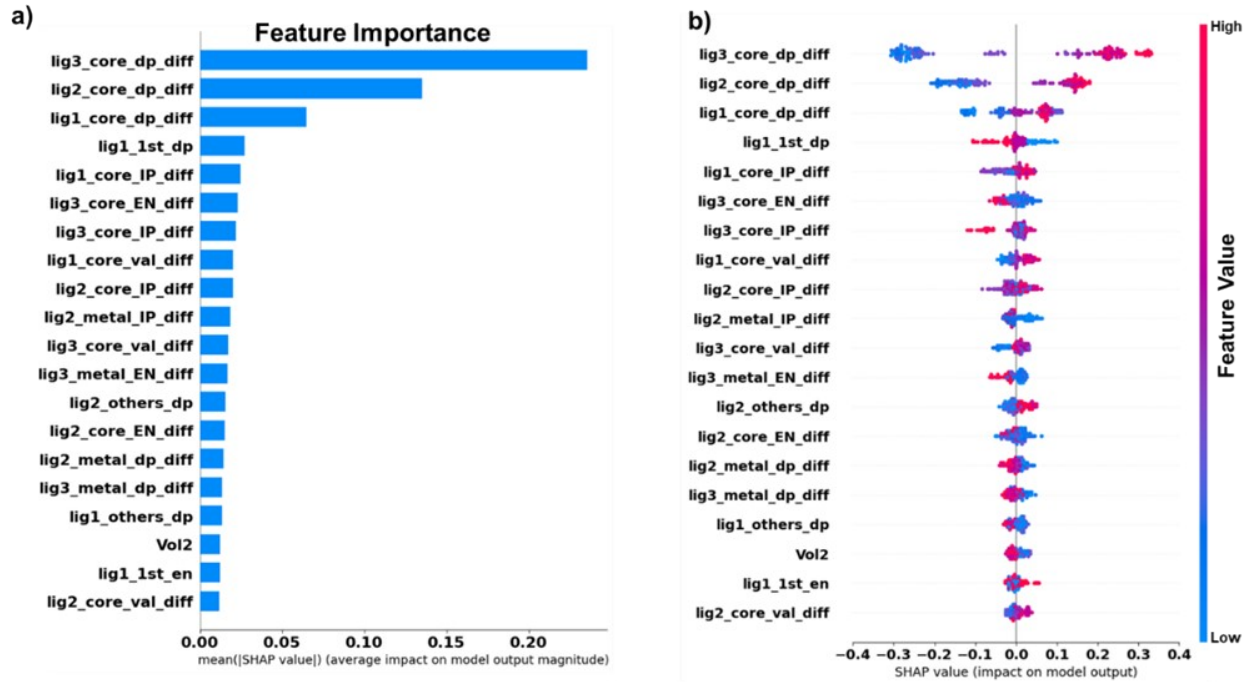

Figure S13. For $^{\Delta E}{}_{*CO \rightarrow *COH}$ prediction, SHAP plot demonstrating the a) feature importance of higher contributing features among others, and b) bee-swarm plot showing the distribution of feature values with their corresponding SHAP value for most contributing features.

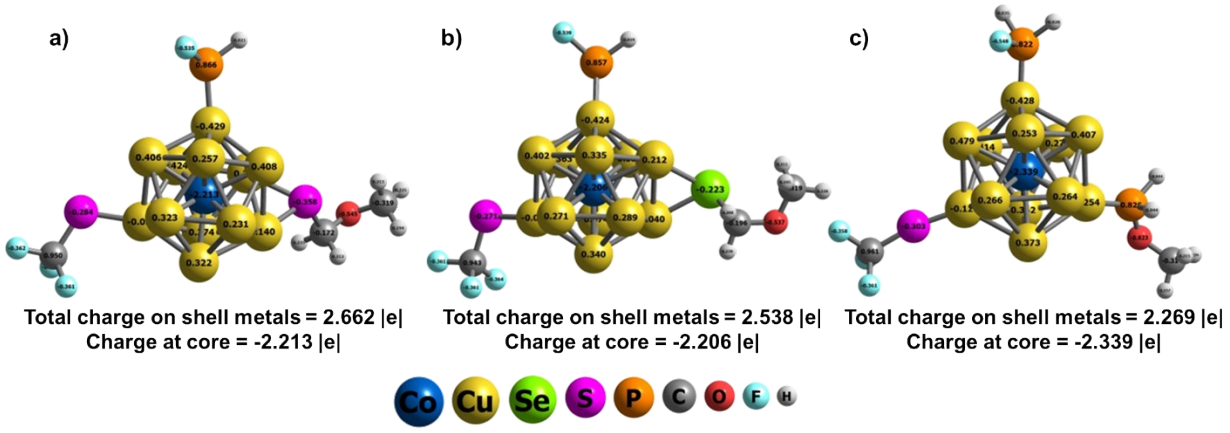

Figure S14. Charge distribution shown on surface and core metals of three selected NCs with a) PH_2F, SCF_3 and SCH_2OCH_3, b) PH_2F, SCF_3 and $SeCH_2OCH_3$ and c) PH_2F, SCF_3 and PH_2OCH_3 as protecting ligands through NBO charge analysis.

Table S13: Comparison of DFT calculated and ML predicted E_{*CO} for some of the randomly selected NC catalysts.

Core Metals	Ligands	Adsorption site	DFT calculated adsorption energy (eV)	Predicted adsorption energy (eV)
Ni	PH_2OCH_3, SCH_3, $SeCH_3$	Between SCH_3 and $SeCH_3$	-1.32	-1.3
Co	PH_2F, SCF_3, SCH_2OCH_3	Between SCF_3 and SCH_2OCH_3	-1.37	-1.3
Fe	PH_2F, PH_3, $SeCH_2OCH_3$	Between PH_3 and $SeCH_2OCH_3$	-1.25	-1.35
Zn	SCH_3, SCH_3, $SeCH_2OCH_3$	Between SCH_3 and SCH_3	-1.31	-1.31
Cu	SCH_2OCH_3, SCH_3, $SeCH_3$	Between SCH_3 and $SeCH_3$	-1.2	-1.36

Table S14: Comparison of DFT calculated and ML predicted E_{*HCO} for some of the randomly selected NC catalysts.

Core Metals	Ligands	Adsorption site	DFT calculated adsorption energy (eV)	Predicted adsorption energy (eV)
Co	PH_2F, SCF_3, SCH_2OCH_3	Between SCF_3 and SCH_2OCH_3	-2.46	-2.35
Ni	PH_2OCH_3, SCH_3, $SeCH_3$,	Between $SeCH_2OCH_3$ and $SeCH_3$	-2.41	-2.21
Fe	PH_2F, PH_3, $SeCH_2OCH_3$,	Between PH_3 and $SeCH_2OCH_3$	-2.19	-2.39
Mn	PH_2F, PH_3, SCH_2OCH_3	Between PH_3 and SCH_2OCH_3	-1.93	-1.84
Cu	PH_2OCH_3, SCF_3, $SeCH_3$	Between SCF_3 and $SeCH_3$	-2.22	-2.19

Table S15: Comparison of DFT calculated and ML predicted E_{*COH} for some of the randomly selected NC catalysts.

Core Metals	Ligands	Adsorption site	DFT calculated adsorption energy (eV)	Predicted adsorption energy (eV)
Co	PH_2F, SCF_3, $SeCH_2OCH_3$	Between PH_2F and SCF_3	-3.05	-3.02
Ni	SCF_3, SCH_3, $SeCF_3$	Between SCH_3 and $SeCF_3$	-3.39	-3.45
Fe	PH_2F, SCH_3, $SeCH_3$	Between SCH_3 and $SeCH_3$	-2.79	-2.72
Mn	PH_2F, PH_3, SCH_2OCH_3	Between PH_3 and SCH_2OCH_3	-2.39	-2.38
Cr	PH_2OCH_3, PH_3, $SeCF_3$	Between PH_3 and $SeCF_3$	-2.27	-2.41

Table S16: Comparison of DFT calculated and ML predicted $\Delta E_{*CO \rightarrow *HCO}$ for some of the randomly selected NC catalysts.

Core Metals	Ligands	Adsorption site	DFT calculated reaction energy (eV)	Predicted reaction energy (eV)
Co	PH_2OCH_3, SCH_2OCH_3, SCH_2OCH_3	Between PH_2OCH_3 and SCH_2OCH_3	-0.03	0.03
Ni	SCH_2OCH_3, $SeCH_2OCH_3$, $SeCH_3$	Between $SeCH_2OCH_3$ and $SeCH_3$	0.54	0.49
Fe	PH_3, $SeCH_2OCH_3$, $SeCH_3$	Between $SeCH_2OCH_3$ and $SeCH_3$	0.54	0.62
Zn	PH_2F, PH_3, $SeCF_3$	Between PH_2F and $SeCF_3$	0.40	0.45
Cu	PH_2F, PH_2OCH_3, PH_3	Between PH_2F and PH_2OCH_3	0.38	0.37

Table S17: Comparison of DFT calculated and ML predicted $\Delta E_{*CO \rightarrow *COH}$ for some of the randomly selected NC catalysts.

Core Metals	Ligands	Adsorption site	DFT calculated reaction energy (eV)	Predicted reaction energy (eV)
Co	$SeCF_3$, $SeCH_2OCH_3$, $SeCH_3$	Between $SeCH_2OCH_3$ and $SeCH_3$	1.54	1.44
Ni	SCH_3, SCH_3, $SeCH_3$	Between SCH_3 and SCH_3	1.25	1.21
Fe	PH_3, $SeCH_2OCH_3$, $SeCH_3$	Between $SeCH_2OCH_3$ and $SeCH_3$	1.93	1.73
Mn	PH_2F, SCH_3, $SeCH_2OCH_3$	Between SCH_3 and $SeCH_2OCH_3$	2.2	2.06
Cu	PH_2OCH_3, $SeCF_3$, $SeCH_2OCH_3$	Between PH_2OCH_3 and $SeCF_3$	1.21	1.01

Table S18: Charge at core in 3d metal doped $Cu_{12}X$ core shell bare nanocluster

Core Metal	Charge at core (q)
Ti	-4.25
V	-3.67
Cr	-3.40
Mn	-2.97
Fe	-2.78
Co	-2.41
Ni	-2.23
Cu	-1.85
Zn	-1.08

Table S19: The charge at core metal (Q_{core},$Q_{core}^{'}$), and on the shell metal (Q_1,$Q_1^{'}$) belongs to adsorption site along with their corresponding *CO adsorption energy and reaction energy change for $*CO \rightarrow *HCO$. Q_{core},$Q_{core}^{'}$ indicate charge at core metal before and after *CO adsorption whereas Q_1,$Q_1^{'}$ indicate charge at shell metal before and after *CO adsorption.

Core Metal	Nearest Ligands	3rd Ligand	Q_1	$Q_1^{'}$	Q_{core}	$Q_{core}^{'}$	E_{*CO} (eV)	$\Delta E_{*CO \rightarrow *HCO}$ (eV)
Mn	PH_2F and PH_2OCH_3	$SeCF_3$	0.38	0.02	-2.99	-2.87	-1.44	0.90
Mn	PH_2F and $SeCF_3$	PH_2OCH_3	0.37	-0.01	-2.99	-2.93	-1.28	0.91
Fe	$SeCF_3$ and PH_2F	PH_2F	0.43	-0.09	-2.64	-2.33	-1.35	0.42
Fe	PH_2F and PH_2F	$SeCF_3$	0.31	-0.14	-2.64	-2.49	-1.43	0.49
Co	SCF_3 and SCH_2OCH_3	$SeCH_3$	0.47	0.11	-2.55	-2.47	-1.15	0.59
Co	SCH_3 and $SeCH_3$	PH_2OCH_3	0.38	-0.06	-2.21	-2.13	-1.36	0.22
Co	SCF_3 and SCH_2OCH_3	PH_2F	0.32	-0.05	-2.21	-2.14	-1.37	0.27
Ni	PH_2F and SCF_3	PH_3	0.36	-0.10	-2.12	-2.11	-1.46	0.15
Ni	PH_3 and SCF_3	PH_2F	0.34	-0.20	-2.12	-2.09	-1.38	0.28
Ni	SCH_3 and $SeCH_3$	PH_2OCH_3	0.24	-0.15	-2.05	-2.04	-1.33	0.27

Table S20: The charge at core metal (Q_{core},$Q_{core}^{'}$), and on the shell metals (Q_1,Q_2,$Q_1^{'}$, $Q_2^{'}$) belongs to adsorption site along with their corresponding *HCO adsorption energy and reaction energy change for $*CO \rightarrow$ *HCO. Q_{core},$Q_{core}^{'}$ indicate charge at core metal before and after *HCO adsorption whereas Q_1,Q_2 indicate charge at shell metals before *HCO adsorption and $Q_1^{'}$, $Q_2^{'}$ indicate charge at shell metals after *HCO adsorption.

Core Metal	Nearest Ligands	3rd Ligand	Q_1	$Q_1^{'}$	Q_2	$Q_2^{'}$	Q_{core}	$Q_{core}^{'}$	E_{*HCO} (eV	$\Delta E_{*CO \rightarrow *HCO}$ (eV)
Mn	PH_2F and PH_2OCH_3	$SeCF_3$	0.38	0.27	0.39	0.39	-2.99	-2.46	-1.91	0.90
Mn	PH_2F and $SeCF_3$	PH_2OCH_3	0.40	0.36	0.37	0.24	-2.99	-2.48	-1.73	0.91
Fe	$SeCF_3$ and PH_2F	PH_2F	0.42	0.24	0.43	0.42	-2.64	-2.77	-2.29	0.42
Fe	PH_2F and PH_2F	$SeCF_3$	0.31	0.28	0.39	0.40	-2.64	-2.81	-2.31	0.49
Co	SCF_3 and SCH_2OCH_3	$SeCH_3$	0.47	0.33	0.47	0.41	-2.55	-2.16	-2.05	0.59
Co	SCH_3 and $SeCH_3$	PH_2OCH_3	0.38	0.26	0.33	0.37	-2.21	-2.43	-2.51	0.22
Co	SCF_3 and SCH_2OCH_3	PH_2F	0.37	0.37	0.32	0.28	-2.21	-2.44	-2.46	0.27
Ni	PH_2F and SCF_3	PH_3	0.36	0.30	0.34	0.32	-2.12	-2.09	-2.68	0.15
Ni	PH_3 and SCF_3	PH_2F	0.34	0.42	0.33	0.30	-2.12	-2.14	-2.47	0.28
Ni	SCH_3 and $SeCH_3$	PH_2OCH_3	0.28	0.23	0.24	0.33	-2.05	-2.12	-2.42	0.27

Table S21: The total charge of CO ($Q_{CO}^{'}$) and HCO ($Q_{HCO}^{'}$), after *CO and *HCO adsorption, respectively.

Core Metal	Nearest Ligands	3rd Ligand	$Q_{CO}^{'}$	$Q_{HCO}^{'}$
Mn	PH_2F and PH_2OCH_3	$SeCF_3$	0.11	-0.42
Mn	PH_2F and $SeCF_3$	PH_2OCH_3	0.15	-0.34
Fe	$SeCF_3$ and PH_2F	PH_2F	0.15	-0.36
Fe	PH_2F and PH_2F	$SeCF_3$	0.10	-0.44
Co	SCF_3 and SCH_2OCH_3	$SeCH_3$	0.14	-0.35
Co	SCH_3 and $SeCH_3$	PH_2OCH_3	0.12	-0.39
Co	SCF_3 and SCH_2OCH_3	PH_2F	0.14	-0.37
Ni	PH_2F and SCF_3	PH_3	0.15	-0.40
Ni	PH_3 and SCF_3	PH_2F	0.13	-0.40
Ni	SCH_3 and $SeCH_3$	PH_2OCH_3	0.11	-0.38

Data and codes

All the data and codes can be found in the following github link:

https://github.com/dips96/Conformal-active-learning

References

1 C. Molnar, Introduction To Conformal Prediction With Python, https://christophmolnar.com/books/conformal-prediction/, (accessed 28 February 2024).