Supplementary Information

Machine Learning-based Screening of Mn-PNP Catalysts for CO_2 Reduction Reaction Using Region-wise Ligand-encoded Feature Matrix

Amitabha Das,[†‡] Diptendu Roy,[†‡] Shyama Charan Mandal,[‡±] and Biswarup Pathak[‡*]

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

[±]Present address: SUNCAT Center for Interface Science and Catalysis, Department of Chemical Engineering, Stanford University, 443 Via Ortega, Stanford, CA 94305, USA

[±]SUNCAT Center for Interface Science and Catalysis, SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA

[†]A.D. and D.R. have contributed equally.

*Email: biswarup@iiti.ac.in

Contents

A. Gaussian full references

Gaussian 09, Revision D.01, M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery, Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, O. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian, Inc., Wallingford CT, 2013.

B. DFT calculated descriptors for all the 44 catalysts.

Table S1. DFT calculated ΔΔG‡ and ΔG‡ values for all the 44 catalysts. All the values are in eV.

SL No.	L_1	L_2	L_3	$\Delta\Delta G^{\ddagger RN}$	$\Delta\Delta G^{\ddagger N}$	$\Delta\Delta G^{\ddagger M}$	$\Delta G^{\ddagger TS1}$	$\Delta G^{\ddagger TS2}$	$\Delta G^{\ddagger TS3}$
1	**Me**	**Me**	**H**	0.58	1.70	2.04	0.35	1.04	-0.01
2	**Ph**	**Ph**	**H**	0.65	1.75	2.08	0.41	1.05	-0.08
3	**Ph**	**H**	**H**	0.53	1.86	2.19	0.18	1.17	0.07
4	**Pr**	**Me**	**H**	0.57	1.71	2.04	0.36	0.85	-0.09
5	**^{t}Bu**	**Et**	**H**	0.46	1.60	1.93	0.24	0.98	0.06
6	**^{i}Pr**	**OMe**	**H**	0.82	1.84	2.26	0.47	1.04	-0.02
7	**Et**	**CF_3**	**H**	1.18	1.72	2.18	0.29	1.05	0.66
8	**H**	**H**	**Me**	0.55	1.79	2.13	0.24	1.08	-0.02
9	**H**	**Pr**	**Me**	0.55	1.76	2.09	0.27	1.05	0.17
10	**H**	**Et**	**Me**	0.56	1.78	2.11	0.11	1.07	0.07
11	**Me**	**^{i}Pr**	**Me**	0.41	1.70	2.03	0.20	0.99	0.01
12	**^{i}Pr**	**Et**	**Me**	0.40	1.68	2.01	0.18	0.96	-0.06
13	**OEt**	**Et**	**Me**	0.64	1.94	2.28	0.41	1.03	-0.07
14	**OEt**	**^{i}Pr**	**Me**	1.01	1.97	2.30	0.15	1.06	0.00
15	**CF_3**	**^{i}Pr**	**Me**	0.70	1.89	2.23	0.36	1.20	0.15
16	**Ph**	**Ph**	**Et**	0.52	1.71	2.05	0.26	1.03	-0.08
17	**H**	**^{t}Bu**	**Et**	0.75	1.63	1.97	0.51	0.91	-0.03
18	**OMe**	**Me**	**Et**	0.55	1.83	2.16	0.22	1.14	0.05
19	**Pr**	**Et**	**Et**	0.47	1.65	1.98	0.25	0.96	0.02
20	**OMe**	**Pr**	**Et**	0.63	1.73	2.15	0.25	0.92	0.00
21	**Ph**	**CF_3**	**Et**	1.06	1.77	2.10	0.43	1.14	0.54
22	**CF_3**	**Pr**	**Et**	0.98	1.90	2.24	0.44	1.21	0.49
23	**^{i}Pr**	**H**	**Pr**	0.54	1.77	2.10	0.33	1.04	-0.04
24	**OEt**	**Me**	**Pr**	0.56	1.83	2.16	0.25	1.07	-0.03
25	**Et**	**Ph**	**Pr**	0.53	1.64	1.97	0.30	0.99	-0.10
26	**OMe**	**Et**	**Pr**	0.59	1.86	2.19	0.29	1.05	0.01
27	**OEt**	**Pr**	**Pr**	0.72	2.16	2.14	0.34	0.97	-0.06
28	**CF_3**	**H**	**Pr**	0.93	1.93	2.26	0.51	1.16	0.14
29	**CF_3**	**OEt**	**Pr**	0.80	2.15	2.49	0.49	1.23	0.35
30	**^{t}Bu**	**^{t}Bu**	**^{i}Pr**	0.94	1.43	1.76	0.68	0.61	0.35
31	**^{i}Pr**	**^{i}Pr**	**^{i}Pr**	0.69	1.38	1.71	0.33	0.70	0.27
32	**H**	**OEt**	**^{i}Pr**	0.70	1.99	2.32	0.46	1.12	-0.02
33	**Et**	**Me**	**^{i}Pr**	0.50	1.70	2.03	0.23	1.00	0.12
34	**Ph**	**^{i}Pr**	**^{i}Pr**	0.60	1.54	1.87	0.29	0.94	-0.05
35	**^{t}Bu**	**Pr**	**^{i}Pr**	0.65	1.70	2.03	0.07	1.01	0.38
36	**OMe**	**OEt**	**iPr**	0.59	1.79	2.12	0.32	1.13	-0.01
37	**CF_3**	**H**	**iPr**	0.76	2.06	2.39	0.30	1.30	0.13
38	**OMe**	**OMe**	**^{t}Bu**	0.72	2.01	2.34	0.50	1.20	0.03

39	**Ph**	**H**	**^{t}Bu**	0.56	1.89	2.22	0.33	1.22	-0.01
40	**Me**	**Ph**	**^{t}Bu**	0.43	1.68	2.01	0.20	1.02	0.02
41	**Ph**	**Pr**	**^{t}Bu**	0.42	1.59	1.92	0.20	0.93	-0.02
42	**^{t}Bu**	**OMe**	**^{t}Bu**	0.72	1.92	2.17	0.47	1.03	-0.03
43	**CF_3**	**Me**	**^{t}Bu**	0.76	1.96	2.30	0.27	1.11	0.13
44	**OMe**	**CF_3**	**^{t}Bu**	1.01	1.93	2.26	0.55	1.23	0.35

C. Optimized hyperparameters values.

Table S2. Optimized values of the hyperparameters for the best fitted ML algorithms for all the considered descriptors.

Descriptors	ML Models	Optimized hyperparameters for best fitted models
$\Delta\Delta G^{\ddagger RN}$	**XGBR**	n_estimators: 100, min_child_weight: 3, max_depth: 5, learning_rate: 0.05, booster: gbtree, base_score: 0.5
$\Delta\Delta G^{\ddagger N}$	**GBR**	n_estimators: 100, min_samples_split: 10, min_samples_leaf: 3, max_depth:10, learning_rate: 0.2
$\Delta\Delta G^{\ddagger M}$	**XGBR**	n_estimators: 1500, min_child_weight: 4, learning_rate: 0.1, booster: gbtree, max_depth: 15, base_score: 1
$\Delta G^{\ddagger TS1}$	**XGBR**	n_estimators: 100, min_child_weight: 2, learning_rate: 0.05, booster: gbtree, max_depth: 15, base_score: 0.25
$\Delta G^{\ddagger TS2}$	**GBR**	n_estimators: 100, min_samples_split: 10, min_samples_leaf: 3, max_depth: 10, learning_rate: 0.5
$\Delta G^{\ddagger TS3}$	**GBR**	n_estimators: 900, min_samples_split: 5, min_samples_leaf: 5, max_depth: 3, learning_rate: 0.1

D. Performance analysis of ML models

Table S3. Performance analysis (RMSE) of all the considered ML models for all the descriptors. All the values are in eV.

Algorithm	Error	$\Delta\Delta G^{\ddagger RN}$	$\Delta\Delta G^{\ddagger N}$	$\Delta\Delta G^{\ddagger M}$	$\Delta G^{\ddagger TS1}$	$\Delta G^{\ddagger TS2}$	$\Delta G^{\ddagger TS3}$
Linear Regression	Train (RMSE)	0.08	0.09	0.07	0.09	0.06	0.10
	Test (RMSE)	0.21	0.11	0.13	0.10	0.12	0.16
KRR	Train (RMSE)	0.12	0.09	0.07	0.09	0.01	0.11
	Test (RMSE)	0.15	0.10	0.13	0.09	0.11	0.10
RFR	Train (RMSE)	0.14	0.09	0.07	0.11	0.07	0.13
	Test (RMSE)	0.13	0.13	0.13	0.10	0.11	0.08
GBR	Train (RMSE)	0.11	0.01	0.13×10^{-4}	0.07	0.01	0.01
	Test (RMSE)	0.15	0.09	0.12	0.08	0.10	0.08
XGBR	Train (RMSE)	0.08	0.87×10^{-4}	0.03	0.04	0.02	0.09
	Test (RMSE)	0.14	0.10	0.12	0.07	0.11	0.09

E. Best fitted ML models for TS1, TS2 and TS3 and their performance analysis.

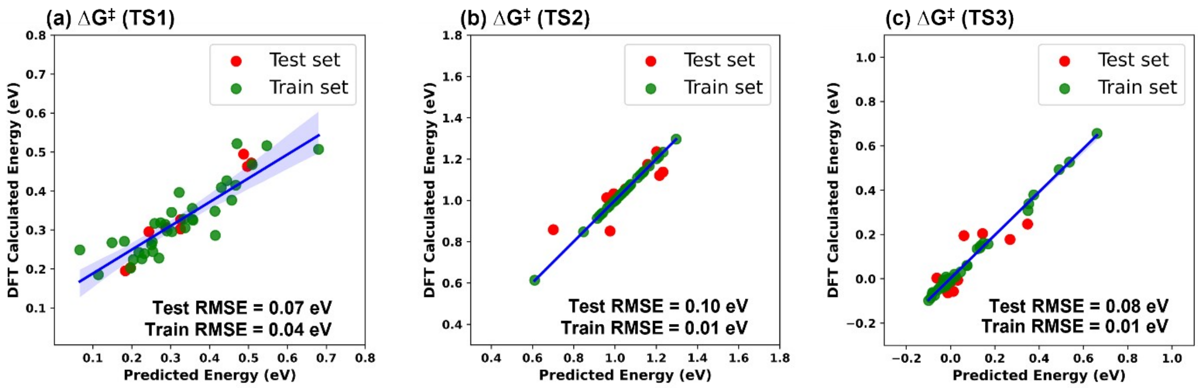

Fig. S1 Plot of DFT calculated vs ML predicted activation barriers of the transition states: (a) TS1, (b) TS2 and (c) TS3. All the energy values are in eV.

F. Cross-Validation (CV) analysis for best fitted machine learning models.

An evaluation method called Cross-Validation is used in machine learning to determine whether the model is stable and efficient enough when trained on the subset of input data sets and tested on the unseen set of datasets. Here, we used the k-fold cross validation method. The data sample is split into 'k' number of smaller samples, and in our case, k is 5 and the data sets is split into 5-fold (CVi, i = 1, 2, 3, 4 and 5) sub-datasets (train and test). First, for example, CV2-5 subsets are considered for training, and the CV1 subset is left for the validation of the trained model. This process is repeated four more times with different subsets for the validation of the trained model. Cross-validation reduces the chances of overfitting and increases the generalizability of the parameters used in the machine learning model. Fig. S2 shows the concept of a 5-fold cross-validation loop. The root-mean-square error (RMSE) of each iteration and the mean of the RMSE for the total 5 iterations are computed.

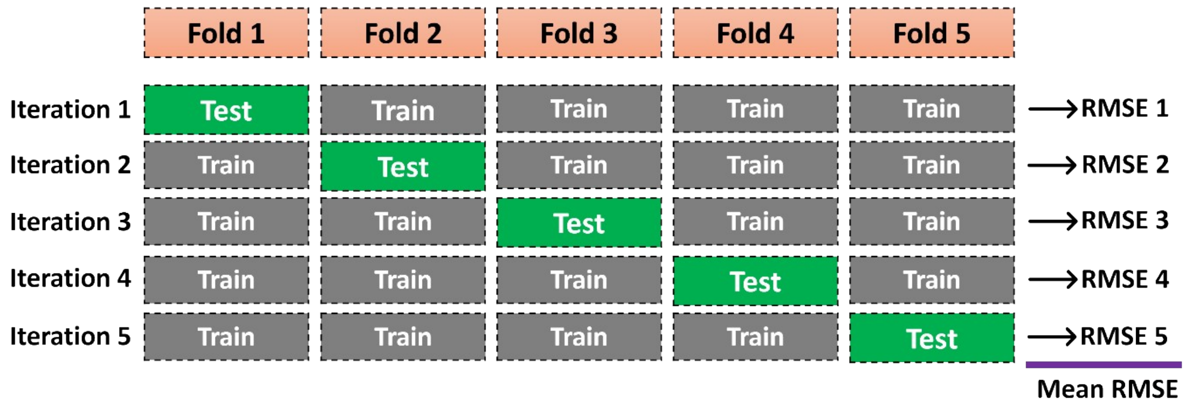

Fig. S2 The scheme explaining of k-fold (k = 5) Cross-Validation (CV) method.

Table S4. Root-Mean-Square Error (RMSE) Score for each fold (CVi) of the 5-Fold CV using best fitted ML models for descriptors. Here, Avg. is the Mean RMSE of all the CV. All the values are in eV.

Data sets	Method	CV1	CV2	CV3	CV4	CV5	Avg.
$\Delta\Delta G^{‡RN}$	**XGBR**	0.14	0.17	0.21	0.15	0.11	0.16
$\Delta\Delta G^{‡N}$	**GBR**	0.10	0.17	0.10	0.11	0.13	0.12
$\Delta\Delta G^{‡M}$	**XGBR**	0.15	0.10	0.16	0.12	0.12	0.13
$\Delta G^{‡TS1}$	**XGBR**	0.10	0.17	0.12	0.10	0.11	0.12

$\Delta G^{\ddagger TS2}$	**GBR**	0.09	0.13	0.11	0.08	0.13	0.11
$\Delta G^{\ddagger TS3}$	**GBR**	0.15	0.10	0.09	0.11	0.15	0.12

G. ML predicted descriptors values for all the 286 catalysts

Table S5. ML predicted $\Delta\Delta G^{\ddagger}$ and $\Delta G^{\ddagger}$ values for all the 286 catalysts. All the values are in eV.

SL No.	L1	L2	L3	$\Delta\Delta G^{\ddagger RN}$	$\Delta\Delta G^{\ddagger N}$	$\Delta\Delta G^{\ddagger M}$	$\Delta G^{\ddagger TS1}$	$\Delta G^{\ddagger TS2}$	$\Delta G^{\ddagger TS3}$
45	**H**	**H**	**H**	0.68	1.82	2.16	0.35	1.13	0.10
46	**Et**	**Et**	**H**	0.57	1.65	2.01	0.26	1.01	0.07
47	**^{t}Bu**	**^{t}Bu**	**H**	0.75	1.66	1.97	0.50	0.93	0.05
48	**Pr**	**Pr**	**H**	0.67	1.66	2.04	0.28	0.92	0.02
49	**^{i}Pr**	**^{i}Pr**	**H**	0.74	1.68	2.01	0.40	0.99	0.02
50	**OMe**	**OMe**	**H**	0.75	1.89	2.28	0.46	1.11	0.03
51	**OEt**	**OEt**	**H**	0.78	1.97	2.29	0.48	1.09	-0.02
52	**Me**	**H**	**H**	0.57	1.80	2.12	0.28	1.11	0.00
53	**Et**	**H**	**H**	0.56	1.81	2.13	0.22	1.11	0.13
54	**^{t}Bu**	**H**	**H**	0.68	1.73	2.06	0.48	0.99	0.08
55	**^{i}Pr**	**H**	**H**	0.67	1.78	2.15	0.33	1.08	0.06
56	**Pr**	**H**	**H**	0.63	1.78	2.12	0.25	1.04	0.10
57	**OMe**	**H**	**H**	0.68	1.89	2.25	0.38	1.13	0.07
58	**OEt**	**H**	**H**	0.71	2.02	2.31	0.38	1.14	0.02
59	**Ph**	**Me**	**H**	0.54	1.71	2.00	0.33	1.04	-0.07
60	**Et**	**Me**	**H**	0.53	1.66	2.04	0.26	1.00	-0.03
61	**^{t}Bu**	**Me**	**H**	0.56	1.64	1.94	0.37	0.91	-0.03
62	**^{i}Pr**	**Me**	**H**	0.56	1.70	2.00	0.33	0.97	-0.03
63	**OMe**	**Me**	**H**	0.57	1.86	2.19	0.35	1.12	-0.02
64	**OEt**	**Me**	**H**	0.58	1.88	2.22	0.36	1.08	-0.07
65	**Et**	**Ph**	**H**	0.57	1.69	2.02	0.28	1.01	-0.06
66	**^{t}Bu**	**Ph**	**H**	0.63	1.68	1.98	0.40	0.96	-0.05
67	**^{i}Pr**	**Ph**	**H**	0.63	1.71	2.03	0.31	1.01	-0.15
68	**Pr**	**Ph**	**H**	0.55	1.61	1.96	0.26	0.94	-0.08
69	**OMe**	**Ph**	**H**	0.63	1.89	2.23	0.37	1.12	-0.08
70	**OEt**	**Ph**	**H**	0.68	1.98	2.26	0.38	1.09	-0.15
71	**^{i}Pr**	**Et**	**H**	0.56	1.66	1.99	0.25	0.97	0.00
72	**Pr**	**Et**	**H**	0.53	1.61	2.02	0.25	0.94	0.03
73	**OMe**	**Et**	**H**	0.57	1.87	2.23	0.25	1.06	0.04
74	**OEt**	**Et**	**H**	0.64	1.96	2.27	0.33	1.05	-0.01
75	**^{i}Pr**	**^{t}Bu**	**H**	0.75	1.66	1.95	0.49	0.94	0.05
76	**Pr**	**^{t}Bu**	**H**	0.67	1.66	2.03	0.25	0.89	0.03
77	**OMe**	**^{t}Bu**	**H**	0.75	1.83	2.16	0.49	0.96	0.01

78	**OEt**	**^{t}Bu**	**H**	0.79	1.94	2.20	0.50	0.97	0.00
79	**Pr**	**^{i}Pr**	**H**	0.67	1.68	2.01	0.27	0.92	-0.02
80	**OEt**	**^{i}Pr**	**H**	0.79	1.97	2.28	0.40	1.05	0.02
81	**OMe**	**Pr**	**H**	0.68	1.80	2.20	0.28	0.99	-0.02
82	**OEt**	**Pr**	**H**	0.70	2.04	2.21	0.30	0.98	-0.06
83	**OEt**	**OMe**	**H**	0.76	1.96	2.30	0.48	1.12	-0.05
84	**CF_3**	**CF_3**	**H**	0.96	1.82	2.23	0.48	1.15	0.56
85	**CF_3**	**H**	**H**	0.91	1.90	2.25	0.43	1.19	0.41
86	**CF_3**	**Me**	**H**	0.80	1.92	2.29	0.38	1.13	0.37
87	**CF_3**	**Ph**	**H**	0.96	1.78	2.15	0.43	1.14	0.57
88	**CF_3**	**^{t}Bu**	**H**	0.96	1.81	2.17	0.50	1.14	0.49
89	**CF_3**	**^{i}Pr**	**H**	0.96	1.85	2.20	0.46	1.15	0.43
90	**CF_3**	**Pr**	**H**	0.91	1.87	2.29	0.34	1.15	0.50
91	**CF_3**	**OMe**	**H**	0.95	1.85	2.23	0.47	1.18	0.54
92	**CF_3**	**OEt**	**H**	0.97	2.03	2.44	0.50	1.18	0.54
93	**Me**	**Me**	**Me**	0.49	1.68	2.04	0.22	1.03	0.00
94	**Ph**	**Ph**	**Me**	0.57	1.70	2.07	0.32	1.05	-0.10
95	**Et**	**Et**	**Me**	0.52	1.66	2.01	0.27	1.00	0.01
96	**^{t}Bu**	**^{t}Bu**	**Me**	0.72	1.65	1.95	0.36	0.93	0.01
97	**Pr**	**Pr**	**Me**	0.59	1.68	2.03	0.22	0.93	0.09
98	**^{i}Pr**	**^{i}Pr**	**Me**	0.59	1.70	2.02	0.18	0.99	-0.02
99	**OMe**	**OMe**	**Me**	0.67	1.91	2.27	0.28	1.11	-0.01
100	**OEt**	**OEt**	**Me**	0.79	1.96	2.29	0.34	1.09	-0.06
101	**Me**	**H**	**Me**	0.49	1.76	2.10	0.20	1.08	-0.01
102	**Ph**	**H**	**Me**	0.53	1.82	2.17	0.16	1.11	-0.04
103	**^{t}Bu**	**H**	**Me**	0.64	1.70	2.02	0.32	0.95	0.00
104	**^{i}Pr**	**H**	**Me**	0.52	1.78	2.14	0.15	1.04	-0.01
105	**OMe**	**H**	**Me**	0.60	1.88	2.23	0.22	1.10	0.00
106	**OEt**	**H**	**Me**	0.72	2.00	2.29	0.21	1.10	-0.05
107	**Ph**	**Me**	**Me**	0.49	1.67	2.01	0.22	1.04	-0.04
108	**Et**	**Me**	**Me**	0.49	1.67	2.06	0.21	1.01	-0.01
109	**^{t}Bu**	**Me**	**Me**	0.53	1.62	1.93	0.25	0.92	0.00
110	**Pr**	**Me**	**Me**	0.48	1.73	2.04	0.22	0.91	0.06
111	**OMe**	**Me**	**Me**	0.50	1.87	2.20	0.22	1.13	0.02
112	**OEt**	**Me**	**Me**	0.59	1.87	2.23	0.24	1.08	-0.04
113	**Et**	**Ph**	**Me**	0.51	1.67	2.02	0.28	1.00	-0.12
114	**^{t}Bu**	**Ph**	**Me**	0.62	1.64	1.97	0.36	0.95	-0.10
115	**^{i}Pr**	**Ph**	**Me**	0.50	1.71	2.05	0.18	1.01	-0.20
116	**Pr**	**Ph**	**Me**	0.50	1.60	1.94	0.22	0.95	-0.02
117	**OMe**	**Ph**	**Me**	0.58	1.89	2.23	0.28	1.12	-0.12
118	**OEt**	**Ph**	**Me**	0.69	1.97	2.27	0.34	1.09	-0.19
119	**^{t}Bu**	**Et**	**Me**	0.57	1.60	1.92	0.26	0.92	0.01
120	**Pr**	**Et**	**Me**	0.48	1.64	2.02	0.22	0.95	0.09
121	**OMe**	**Et**	**Me**	0.53	1.89	2.22	0.22	1.07	-0.01
122	**^{i}Pr**	**^{t}Bu**	**Me**	0.72	1.67	1.96	0.28	0.94	0.02
123	**Pr**	**^{t}Bu**	**Me**	0.64	1.68	2.01	0.19	0.91	0.10

124	**OMe**	**^{t}Bu**	**Me**	0.71	1.85	2.15	0.32	0.97	-0.04
125	**OEt**	**^{t}Bu**	**Me**	0.80	1.93	2.19	0.37	0.97	-0.04
126	**Pr**	**^{i}Pr**	**Me**	0.52	1.71	2.04	0.19	0.93	0.06
127	**OMe**	**^{i}Pr**	**Me**	0.61	1.86	2.26	0.22	1.05	-0.05
128	**OMe**	**Pr**	**Me**	0.60	1.83	2.19	0.22	0.99	0.06
129	**OEt**	**Pr**	**Me**	0.70	2.06	2.22	0.25	0.99	0.02
130	**OEt**	**OMe**	**Me**	0.77	1.95	2.29	0.31	1.13	-0.08
131	**CF_3**	**CF_3**	**Me**	0.93	1.85	2.23	0.41	1.17	0.29
132	**CF_3**	**H**	**Me**	0.88	1.90	2.26	0.32	1.16	0.15
133	**CF_3**	**Me**	**Me**	0.78	1.95	2.29	0.30	1.16	0.13
134	**CF_3**	**Ph**	**Me**	0.93	1.78	2.15	0.42	1.16	0.31
135	**CF_3**	**^{t}Bu**	**Me**	0.97	1.84	2.15	0.44	1.17	0.23
136	**CF_3**	**Et**	**Me**	0.93	1.77	2.19	0.31	1.14	0.37
137	**CF_3**	**Pr**	**Me**	0.89	1.91	2.29	0.30	1.18	0.37
138	**CF_3**	**OMe**	**Me**	0.94	1.88	2.25	0.36	1.19	0.28
139	**CF_3**	**OEt**	**Me**	0.97	2.05	2.46	0.42	1.20	0.29
140	**H**	**H**	**Et**	0.64	1.73	2.08	0.35	1.08	-0.01
141	**Me**	**Me**	**Et**	0.53	1.67	2.00	0.23	1.02	0.02
142	**Et**	**Et**	**Et**	0.53	1.67	1.98	0.23	0.99	0.02
143	**^{t}Bu**	**^{t}Bu**	**Et**	0.75	1.63	1.91	0.48	0.91	0.03
144	**Pr**	**Pr**	**Et**	0.64	1.66	1.98	0.30	0.93	0.04
145	**^{i}Pr**	**^{i}Pr**	**Et**	0.63	1.69	1.95	0.34	0.97	0.01
146	**OMe**	**OMe**	**Et**	0.71	1.82	2.21	0.37	1.11	0.02
147	**OEt**	**OEt**	**Et**	0.77	1.94	2.24	0.41	1.07	-0.04
148	**Me**	**H**	**Et**	0.54	1.73	2.06	0.21	1.09	-0.04
149	**Ph**	**H**	**Et**	0.55	1.78	2.12	0.22	1.11	-0.06
150	**Et**	**H**	**Et**	0.52	1.75	2.07	0.22	1.08	0.01
151	**^{i}Pr**	**H**	**Et**	0.57	1.74	2.07	0.32	1.04	-0.03
152	**Pr**	**H**	**Et**	0.62	1.71	2.05	0.30	1.03	0.07
153	**OMe**	**H**	**Et**	0.65	1.81	2.17	0.36	1.12	-0.03
154	**OEt**	**H**	**Et**	0.72	1.97	2.23	0.37	1.10	-0.09
155	**Ph**	**Me**	**Et**	0.51	1.66	2.00	0.22	1.03	-0.02
156	**Et**	**Me**	**Et**	0.50	1.66	2.02	0.23	1.00	-0.01
157	**^{t}Bu**	**Me**	**Et**	0.58	1.60	1.90	0.30	0.91	0.02
158	**Pr**	**Me**	**Et**	0.54	1.70	2.00	0.25	0.91	0.01
159	**^{i}Pr**	**Me**	**Et**	0.46	1.69	1.98	0.24	0.97	0.04
160	**OEt**	**Me**	**Et**	0.59	1.89	2.19	0.26	1.07	-0.02
161	**Et**	**Ph**	**Et**	0.53	1.68	2.01	0.24	1.00	-0.11
162	**^{t}Bu**	**Ph**	**Et**	0.63	1.63	1.93	0.38	0.93	-0.07
163	**^{i}Pr**	**Ph**	**Et**	0.51	1.71	2.00	0.26	0.99	-0.16
164	**Pr**	**Ph**	**Et**	0.52	1.59	1.92	0.28	0.95	-0.07
165	**OMe**	**Ph**	**Et**	0.59	1.81	2.19	0.27	1.12	-0.09
166	**OEt**	**Ph**	**Et**	0.67	1.97	2.22	0.31	1.07	-0.17
167	**^{t}Bu**	**Et**	**Et**	0.58	1.60	1.90	0.28	0.92	0.01
168	**^{i}Pr**	**Et**	**Et**	0.45	1.68	1.95	0.25	0.96	-0.04
169	**OMe**	**Et**	**Et**	0.54	1.82	2.18	0.23	1.06	0.00

170	**OEt**	**Et**	**Et**	0.64	1.94	2.23	0.29	1.04	-0.06
171	**^{i}Pr**	**^{t}Bu**	**Et**	0.75	1.66	1.90	0.47	0.91	0.05
172	**Pr**	**^{t}Bu**	**Et**	0.69	1.65	1.98	0.29	0.90	0.05
173	**OMe**	**^{t}Bu**	**Et**	0.75	1.76	2.09	0.47	0.96	-0.01
174	**OEt**	**^{t}Bu**	**Et**	0.79	1.91	2.15	0.50	0.95	-0.02
175	**Pr**	**^{i}Pr**	**Et**	0.57	1.70	1.97	0.30	0.92	0.01
176	**OMe**	**^{i}Pr**	**Et**	0.64	1.79	2.19	0.36	1.05	-0.02
177	**OEt**	**^{i}Pr**	**Et**	0.78	1.94	2.22	0.35	1.03	0.01
178	**OEt**	**Pr**	**Et**	0.70	2.01	2.17	0.32	0.98	-0.04
179	**OEt**	**OMe**	**Et**	0.75	1.93	2.23	0.41	1.12	-0.06
180	**CF_3**	**CF_3**	**Et**	0.93	1.84	2.18	0.47	1.15	0.49
181	**CF_3**	**H**	**Et**	0.89	1.89	2.20	0.44	1.17	0.30
182	**CF_3**	**Me**	**Et**	0.78	1.94	2.24	0.32	1.15	0.32
183	**CF_3**	**^{t}Bu**	**Et**	0.97	1.83	2.10	0.51	1.15	0.43
184	**CF_3**	**Et**	**Et**	0.93	1.78	2.14	0.30	1.14	0.52
185	**CF_3**	**^{i}Pr**	**Et**	0.85	1.88	2.16	0.45	1.15	0.36
186	**CF_3**	**OMe**	**Et**	0.94	1.86	2.18	0.47	1.19	0.48
187	**CF_3**	**OEt**	**Et**	0.98	2.04	2.40	0.49	1.18	0.48
188	**H**	**H**	**Pr**	0.63	1.76	2.09	0.38	1.11	-0.01
189	**Me**	**Me**	**Pr**	0.53	1.64	2.00	0.24	1.02	0.02
190	**Ph**	**Ph**	**Pr**	0.57	1.68	2.03	0.33	1.04	-0.08
191	**Et**	**Et**	**Pr**	0.51	1.69	1.97	0.28	0.99	0.02
192	**^{t}Bu**	**^{t}Bu**	**Pr**	0.75	1.66	1.93	0.48	0.92	0.02
193	**Pr**	**Pr**	**Pr**	0.64	1.76	1.97	0.27	0.92	0.03
194	**^{i}Pr**	**^{i}Pr**	**Pr**	0.62	1.69	1.97	0.36	0.98	-0.01
195	**OMe**	**OMe**	**Pr**	0.70	1.87	2.24	0.41	1.11	0.02
196	**OEt**	**OEt**	**Pr**	0.75	1.97	2.23	0.43	1.09	-0.05
197	**Me**	**H**	**Pr**	0.54	1.72	2.06	0.24	1.10	-0.03
198	**Ph**	**H**	**Pr**	0.54	1.79	2.12	0.25	1.13	-0.06
199	**Et**	**H**	**Pr**	0.51	1.77	2.07	0.28	1.09	0.01
200	**^{t}Bu**	**H**	**Pr**	0.69	1.71	2.01	0.47	0.97	-0.03
201	**Pr**	**H**	**Pr**	0.60	1.81	2.04	0.25	1.04	0.07
202	**OMe**	**H**	**Pr**	0.65	1.85	2.19	0.40	1.13	-0.01
203	**OEt**	**H**	**Pr**	0.69	1.99	2.23	0.39	1.13	-0.09
204	**Ph**	**Me**	**Pr**	0.51	1.64	1.97	0.23	1.04	-0.01
205	**Et**	**Me**	**Pr**	0.50	1.63	2.01	0.26	1.00	0.01
206	**^{t}Bu**	**Me**	**Pr**	0.58	1.60	1.91	0.29	0.91	0.02
207	**Pr**	**Me**	**Pr**	0.53	1.74	1.99	0.25	0.91	0.01
208	**^{i}Pr**	**Me**	**Pr**	0.46	1.67	1.98	0.25	0.97	0.03
209	**OMe**	**Me**	**Pr**	0.54	1.82	2.16	0.24	1.12	0.04
210	**^{t}Bu**	**Ph**	**Pr**	0.62	1.64	1.94	0.39	0.95	-0.08
211	**^{i}Pr**	**Ph**	**Pr**	0.50	1.70	2.00	0.28	1.00	-0.17
212	**Pr**	**Ph**	**Pr**	0.52	1.69	1.90	0.24	0.95	-0.07
213	**OMe**	**Ph**	**Pr**	0.59	1.86	2.19	0.33	1.12	-0.08
214	**OEt**	**Ph**	**Pr**	0.64	1.96	2.21	0.35	1.09	-0.17
215	**^{t}Bu**	**Et**	**Pr**	0.56	1.62	1.90	0.29	0.91	0.02

216	**^{i}Pr**	**Et**	**Pr**	0.44	1.68	1.95	0.27	0.96	-0.04
217	**Pr**	**Et**	**Pr**	0.50	1.72	1.97	0.27	0.95	0.03
218	**OEt**	**Et**	**Pr**	0.60	1.97	2.21	0.32	1.04	-0.06
219	**^{i}Pr**	**^{t}Bu**	**Pr**	0.74	1.68	1.92	0.46	0.93	0.03
220	**Pr**	**^{t}Bu**	**Pr**	0.69	1.75	1.96	0.24	0.90	0.04
221	**OMe**	**^{t}Bu**	**Pr**	0.75	1.85	2.13	0.47	0.96	-0.01
222	**OEt**	**^{t}Bu**	**Pr**	0.77	1.96	2.15	0.49	0.96	-0.03
223	**Pr**	**^{i}Pr**	**Pr**	0.56	1.77	1.97	0.26	0.92	0.00
224	**OMe**	**^{i}Pr**	**Pr**	0.64	1.84	2.21	0.40	1.05	-0.02
225	**OEt**	**^{i}Pr**	**Pr**	0.75	1.97	2.22	0.37	1.05	-0.01
226	**OMe**	**Pr**	**Pr**	0.65	1.90	2.14	0.27	0.99	0.01
227	**OEt**	**OMe**	**Pr**	0.74	1.94	2.25	0.44	1.12	-0.06
228	**CF_3**	**CF_3**	**Pr**	0.89	1.94	2.26	0.48	1.16	0.35
229	**CF_3**	**Me**	**Pr**	0.78	1.95	2.30	0.32	1.15	0.20
230	**CF_3**	**Ph**	**Pr**	0.89	1.86	2.17	0.43	1.16	0.37
231	**CF_3**	**^{t}Bu**	**Pr**	0.93	1.93	2.18	0.51	1.16	0.28
232	**CF_3**	**Et**	**Pr**	0.89	1.87	2.22	0.32	1.13	0.43
233	**CF_3**	**^{i}Pr**	**Pr**	0.81	1.96	2.23	0.46	1.17	0.22
234	**CF_3**	**Pr**	**Pr**	0.89	1.98	2.29	0.34	1.17	0.35
235	**CF_3**	**OMe**	**Pr**	0.90	1.96	2.27	0.47	1.19	0.35
236	**H**	**H**	**^{i}Pr**	0.62	1.68	2.13	0.34	1.10	0.11
237	**Me**	**Me**	**^{i}Pr**	0.51	1.59	1.97	0.21	0.89	0.17
238	**Ph**	**Ph**	**^{i}Pr**	0.56	1.56	1.98	0.28	0.92	0.01
239	**Et**	**Et**	**^{i}Pr**	0.51	1.58	1.91	0.26	0.92	0.19
240	**Pr**	**Pr**	**^{i}Pr**	0.62	1.69	2.04	0.21	0.87	0.25
241	**OMe**	**OMe**	**^{i}Pr**	0.68	1.50	2.01	0.30	1.06	0.12
242	**OEt**	**OEt**	**^{i}Pr**	0.73	1.83	2.16	0.40	0.92	0.09
243	**Me**	**H**	**^{i}Pr**	0.51	1.68	2.16	0.23	1.10	0.04
244	**Ph**	**H**	**^{i}Pr**	0.52	1.72	2.17	0.24	1.14	-0.03
245	**Et**	**H**	**^{i}Pr**	0.50	1.76	2.15	0.25	1.13	0.11
246	**^{t}Bu**	**H**	**^{i}Pr**	0.71	1.60	2.03	0.49	0.94	0.23
247	**^{i}Pr**	**H**	**^{i}Pr**	0.61	1.61	2.03	0.34	1.04	0.13
248	**Pr**	**H**	**^{i}Pr**	0.58	1.76	2.17	0.22	1.05	0.21
249	**OMe**	**H**	**^{i}Pr**	0.61	1.63	2.14	0.34	1.12	0.02
250	**Ph**	**Me**	**^{i}Pr**	0.48	1.64	2.03	0.20	0.94	0.04
251	**^{t}Bu**	**Me**	**^{i}Pr**	0.59	1.53	1.85	0.31	0.81	0.29
252	**Pr**	**Me**	**^{i}Pr**	0.51	1.75	2.08	0.22	0.87	0.18
253	**^{i}Pr**	**Me**	**^{i}Pr**	0.51	1.55	1.85	0.25	0.82	0.23
254	**OMe**	**Me**	**^{i}Pr**	0.50	1.60	2.06	0.21	1.08	0.13
255	**OEt**	**Me**	**^{i}Pr**	0.53	1.87	2.20	0.25	0.93	0.08
256	**Et**	**Ph**	**^{i}Pr**	0.51	1.61	2.00	0.26	0.92	-0.03
257	**^{t}Bu**	**Ph**	**^{i}Pr**	0.65	1.50	1.87	0.44	0.86	0.15
258	**Pr**	**Ph**	**^{i}Pr**	0.49	1.68	2.01	0.21	0.92	0.05
259	**OMe**	**Ph**	**^{i}Pr**	0.56	1.56	2.06	0.23	1.10	-0.07
260	**OEt**	**Ph**	**^{i}Pr**	0.62	1.90	2.21	0.33	0.94	-0.11
261	**^{t}Bu**	**Et**	**^{i}Pr**	0.60	1.52	1.82	0.33	0.86	0.32

262	**^{i}Pr**	**Et**	**^{i}Pr**	0.51	1.52	1.78	0.26	0.86	0.18
263	**Pr**	**Et**	**^{i}Pr**	0.47	1.71	2.06	0.23	0.92	0.22
264	**OMe**	**Et**	**^{i}Pr**	0.50	1.59	2.04	0.22	1.09	0.09
265	**OEt**	**Et**	**^{i}Pr**	0.59	1.90	2.21	0.30	0.93	0.07
266	**^{i}Pr**	**^{t}Bu**	**^{i}Pr**	0.77	1.40	1.68	0.47	0.79	0.36
267	**OMe**	**^{t}Bu**	**^{i}Pr**	0.76	1.45	1.91	0.47	0.90	0.25
268	**OEt**	**^{t}Bu**	**^{i}Pr**	0.81	1.80	2.08	0.51	0.84	0.25
269	**Pr**	**^{i}Pr**	**^{i}Pr**	0.61	1.63	1.95	0.23	0.87	0.27
270	**OMe**	**^{i}Pr**	**^{i}Pr**	0.68	1.44	1.96	0.34	0.99	0.14
271	**OEt**	**^{i}Pr**	**^{i}Pr**	0.73	1.81	2.10	0.37	0.87	0.18
272	**OMe**	**Pr**	**^{i}Pr**	0.61	1.64	2.11	0.21	1.00	0.14
273	**OEt**	**Pr**	**^{i}Pr**	0.65	2.01	2.27	0.24	0.94	0.12
274	**CF_3**	**CF_3**	**^{i}Pr**	0.90	1.89	2.24	0.37	1.12	0.37
275	**CF_3**	**Me**	**^{i}Pr**	0.74	2.00	2.33	0.25	1.12	0.17
276	**CF_3**	**Ph**	**^{i}Pr**	0.90	1.91	2.28	0.32	1.13	0.32
277	**CF_3**	**^{t}Bu**	**^{i}Pr**	0.97	1.89	2.15	0.51	1.12	0.43
278	**CF_3**	**Et**	**^{i}Pr**	0.90	1.89	2.25	0.29	1.12	0.43
279	**CF_3**	**^{i}Pr**	**^{i}Pr**	0.88	1.86	2.16	0.35	1.12	0.29
280	**CF_3**	**Pr**	**^{i}Pr**	0.85	2.00	2.36	0.26	1.16	0.39
281	**CF_3**	**OMe**	**^{i}Pr**	0.89	1.89	2.24	0.32	1.19	0.28
282	**CF_3**	**OEt**	**^{i}Pr**	0.91	2.02	2.38	0.42	1.13	0.32
283	**H**	**H**	**^{t}Bu**	0.64	1.86	2.21	0.43	1.16	0.01
284	**Me**	**Me**	**^{t}Bu**	0.53	1.74	2.05	0.23	1.05	-0.01
285	**Ph**	**Ph**	**^{t}Bu**	0.56	1.75	2.08	0.39	1.07	-0.04
286	**Et**	**Et**	**^{t}Bu**	0.53	1.72	2.06	0.32	1.05	0.00
287	**^{t}Bu**	**^{t}Bu**	**^{t}Bu**	0.76	1.74	1.96	0.53	0.99	0.01
288	**Pr**	**Pr**	**^{t}Bu**	0.63	1.70	2.01	0.27	0.97	0.02
289	**^{i}Pr**	**^{i}Pr**	**^{t}Bu**	0.63	1.78	2.05	0.42	1.04	-0.02
290	**OEt**	**OEt**	**^{t}Bu**	0.78	2.00	2.30	0.51	1.11	-0.07
291	**Me**	**H**	**^{t}Bu**	0.53	1.82	2.15	0.22	1.14	-0.02
292	**Et**	**H**	**^{t}Bu**	0.52	1.84	2.18	0.28	1.15	0.04
293	**^{t}Bu**	**H**	**^{t}Bu**	0.69	1.81	2.07	0.49	1.05	-0.01
294	**^{i}Pr**	**H**	**^{t}Bu**	0.57	1.84	2.19	0.40	1.13	-0.02
295	**Pr**	**H**	**^{t}Bu**	0.60	1.77	2.11	0.25	1.09	0.10
296	**OMe**	**H**	**^{t}Bu**	0.65	1.94	2.32	0.45	1.19	0.01
297	**OEt**	**H**	**^{t}Bu**	0.72	2.04	2.33	0.45	1.16	-0.07
298	**Et**	**Me**	**^{t}Bu**	0.50	1.70	2.06	0.22	1.04	-0.03
299	**^{t}Bu**	**Me**	**^{t}Bu**	0.58	1.70	1.93	0.29	0.97	-0.01
300	**Pr**	**Me**	**^{t}Bu**	0.53	1.74	2.01	0.24	0.95	-0.02
301	**^{i}Pr**	**Me**	**^{t}Bu**	0.46	1.75	2.02	0.23	1.02	0.00
302	**OMe**	**Me**	**^{t}Bu**	0.54	1.92	2.23	0.23	1.16	0.03
303	**OEt**	**Me**	**^{t}Bu**	0.58	1.91	2.22	0.25	1.09	-0.06
304	**Et**	**Ph**	**^{t}Bu**	0.51	1.72	2.04	0.33	1.05	-0.07
305	**^{t}Bu**	**Ph**	**^{t}Bu**	0.61	1.71	1.96	0.40	1.00	-0.04
306	**^{i}Pr**	**Ph**	**^{t}Bu**	0.49	1.76	2.04	0.30	1.05	-0.14
307	**OMe**	**Ph**	**^{t}Bu**	0.57	1.93	2.25	0.36	1.16	-0.05

308	**OEt**	**Ph**	**^{t}Bu**	0.67	1.99	2.26	0.40	1.10	-0.14
309	**^{t}Bu**	**Et**	**^{t}Bu**	0.58	1.68	1.94	0.32	0.99	0.00
310	**^{i}Pr**	**Et**	**^{t}Bu**	0.46	1.73	2.03	0.29	1.03	-0.07
311	**Pr**	**Et**	**^{t}Bu**	0.49	1.67	2.00	0.25	0.98	0.01
312	**OMe**	**Et**	**^{t}Bu**	0.54	1.94	2.29	0.29	1.13	0.00
313	**OEt**	**Et**	**^{t}Bu**	0.64	1.97	2.28	0.37	1.08	-0.08
314	**^{i}Pr**	**^{t}Bu**	**^{t}Bu**	0.75	1.76	1.95	0.48	0.99	0.02
315	**Pr**	**^{t}Bu**	**^{t}Bu**	0.68	1.70	1.99	0.24	0.95	0.03
316	**OEt**	**^{t}Bu**	**^{t}Bu**	0.79	1.97	2.18	0.53	1.02	-0.05
317	**Pr**	**^{i}Pr**	**^{t}Bu**	0.56	1.73	2.01	0.26	0.97	-0.01
318	**OMe**	**^{i}Pr**	**^{t}Bu**	0.65	1.93	2.31	0.45	1.12	-0.03
319	**OEt**	**^{i}Pr**	**^{t}Bu**	0.78	2.00	2.30	0.42	1.08	-0.03
320	**OMe**	**Pr**	**^{t}Bu**	0.64	1.85	2.19	0.27	1.06	0.01
321	**OEt**	**Pr**	**^{t}Bu**	0.69	2.06	2.19	0.29	0.99	-0.06
322	**OEt**	**OMe**	**^{t}Bu**	0.76	1.99	2.32	0.50	1.16	-0.07
323	**CF_3**	**CF_3**	**^{t}Bu**	0.93	1.90	2.25	0.55	1.16	0.34
324	**CF_3**	**H**	**^{t}Bu**	0.88	1.91	2.28	0.50	1.20	0.18
325	**CF_3**	**Ph**	**^{t}Bu**	0.93	1.82	2.15	0.46	1.16	0.39
326	**CF_3**	**tBu**	**^{t}Bu**	0.97	1.90	2.14	0.55	1.16	0.28
327	**CF_3**	**Et**	**^{t}Bu**	0.93	1.81	2.21	0.37	1.14	0.41
328	**CF_3**	**iPr**	**^{t}Bu**	0.85	1.94	2.22	0.49	1.17	0.21
329	**CF_3**	**Pr**	**^{t}Bu**	0.89	1.93	2.27	0.34	1.16	0.35
330	**CF_3**	**OEt**	**^{t}Bu**	0.97	2.05	2.45	0.55	1.19	0.33

H. DFT validation of ML predicted energies of all the descriptors.

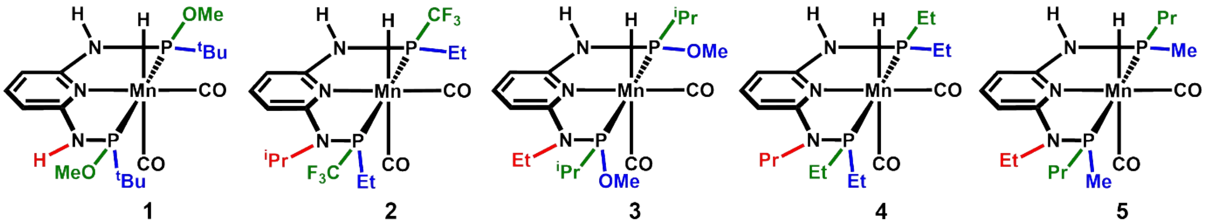

Fig. S3 Considered catalysts for the DFT validation of ML predicted energies of all the descriptors.

Table S6. Comparison of DFT calculated and ML predicted results for all the catalysts mentioned in Fig. S3. All the energies are in eV.

Catalysts	$\Delta G^{\ddagger TS1}$		$\Delta G^{\ddagger TS2}$		$\Delta G^{\ddagger TS3}$	
	DFT	ML	DFT	ML	DFT	ML
1	0.39	0.49	1.03	0.96	-0.01	0.01
2	0.44	0.29	1.13	1.12	0.38	0.43
3	0.46	0.36	0.99	1.05	-0.06	-0.02
4	0.39	0.24	1.02	1.02	-0.02	0.02
5	0.29	0.25	1.01	0.91	-0.01	0.01

I. Proposed best performing catalysts to produce HCOOH from CO_2.

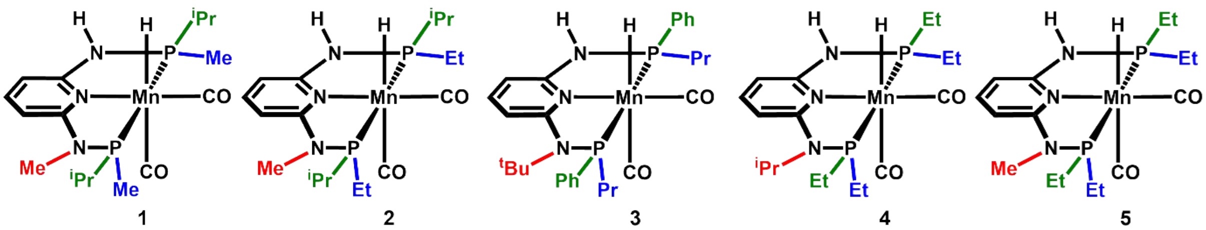

Fig. S4. DFT+ML predicted best catalysts to produce HCOOH.

J. Natural bond orbital (NBO) analysis of catalysts

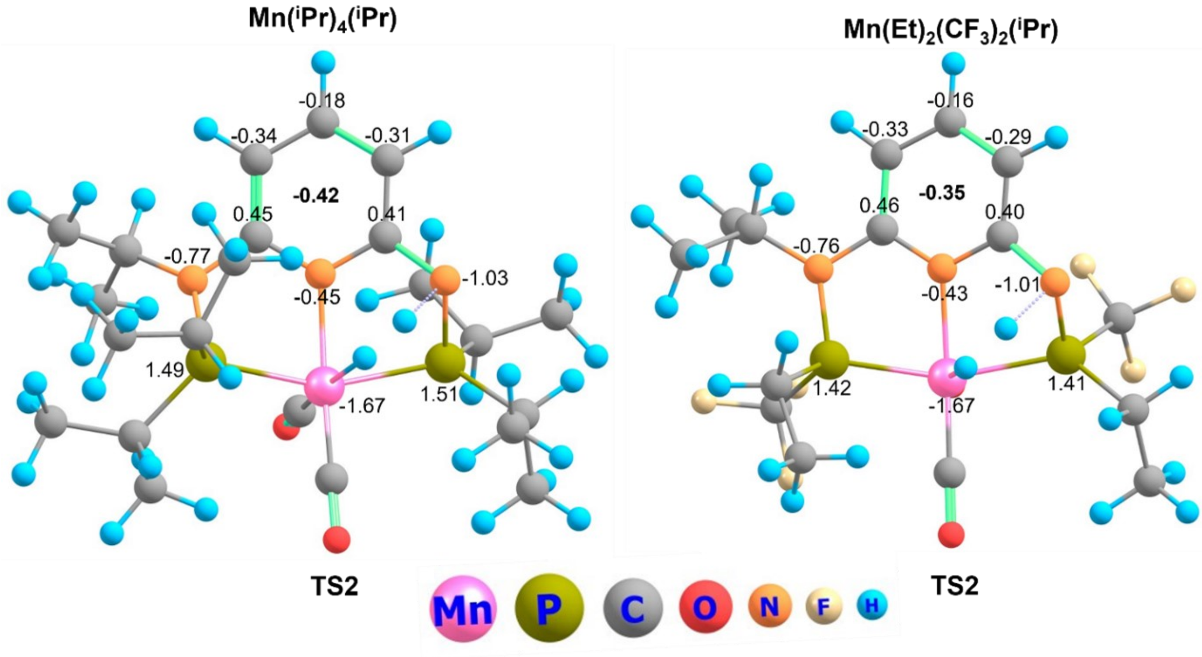

Fig. S5 NBO analysis of catalysts.