Hydrogen adsorption energy trends in Mo/WXY (X, Y= S, Se, Te) regular and Janus TMD monolayers: A first-principles and machine learning study

G. Tejaswini[a], Anjana E Sudheer[a], Amrendra Kumar[b,c], Pavan Kumar Perepu[d],

M. Vallinayagam[e, f], C. Kamal[b,c], S. Mani Prakash[a], M.Posselt[g], M. Zschornak[e, f] and D. Murali[a*]

[a] *Indian Institute of Information Technology Design and Manufacturing, Kurnool, India.*

E-mail: dmurali@iiitk.ac.in

[b] *Theory and Simulations Laboratory, Theoretical and Computational Physics Section,*

Raja Ramanna Centre for Advanced Technology, Indore 452013, India

[c] *Homi Bhabha National Institute, Training School Complex, Mumbai 400094, India*

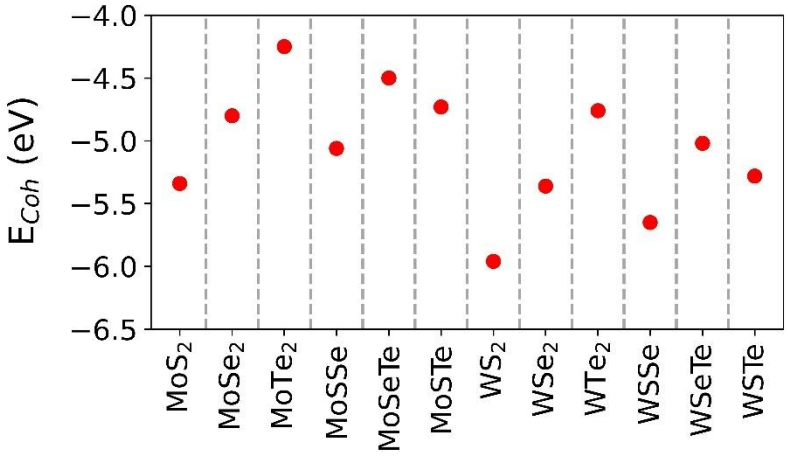
[d] *Indian Institute of Information Technology, Sri City, Andhra Pradesh, India*

[e] *IEP, TU Bergakademie Freiberg, 09599 Freiberg, Germany*

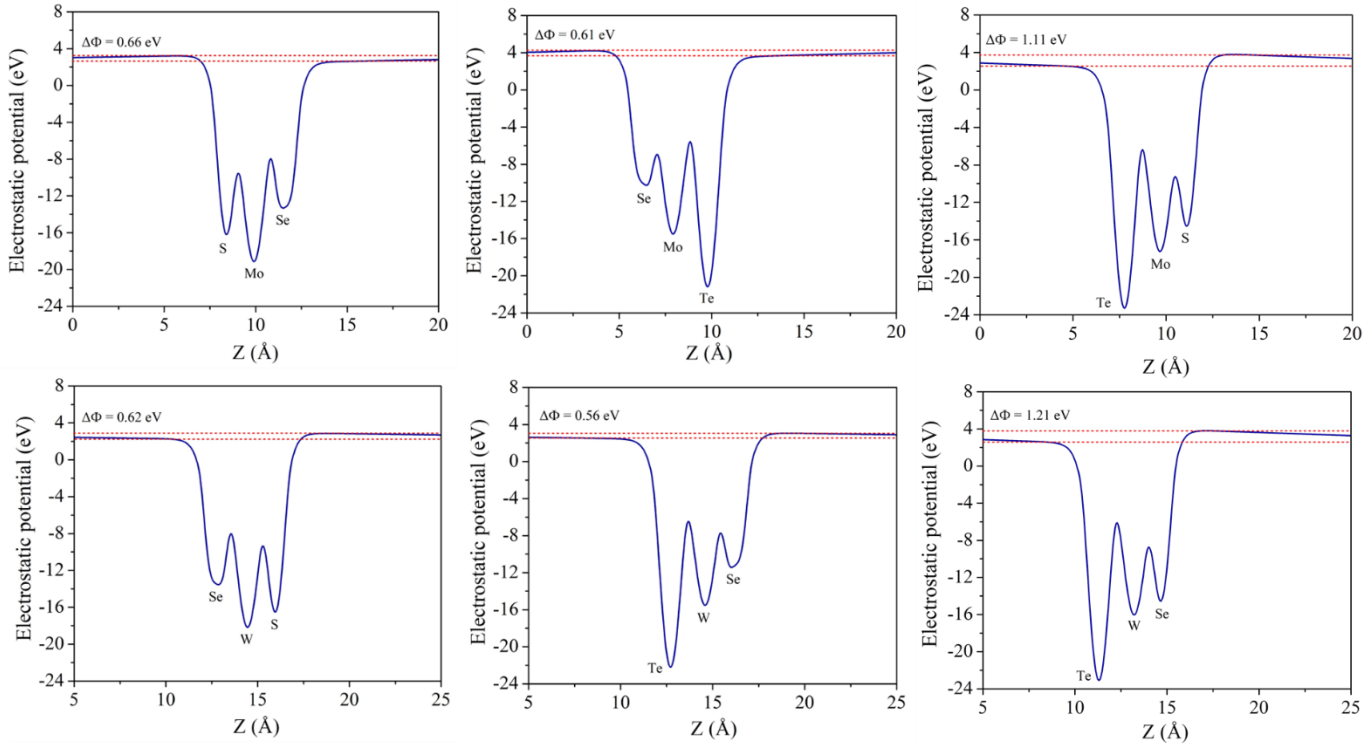
[f] *Fakultät Maschinenbau/Energietechnik/Physik, Hochschule für Technik und Wirtschaft 01069 Dresden, Germany*

[g] *Helmholtz-Zentrum Dresden-Rossendorf, Institute of Ion Beam Physics and Materials Research,01328 Dresden, Germany*

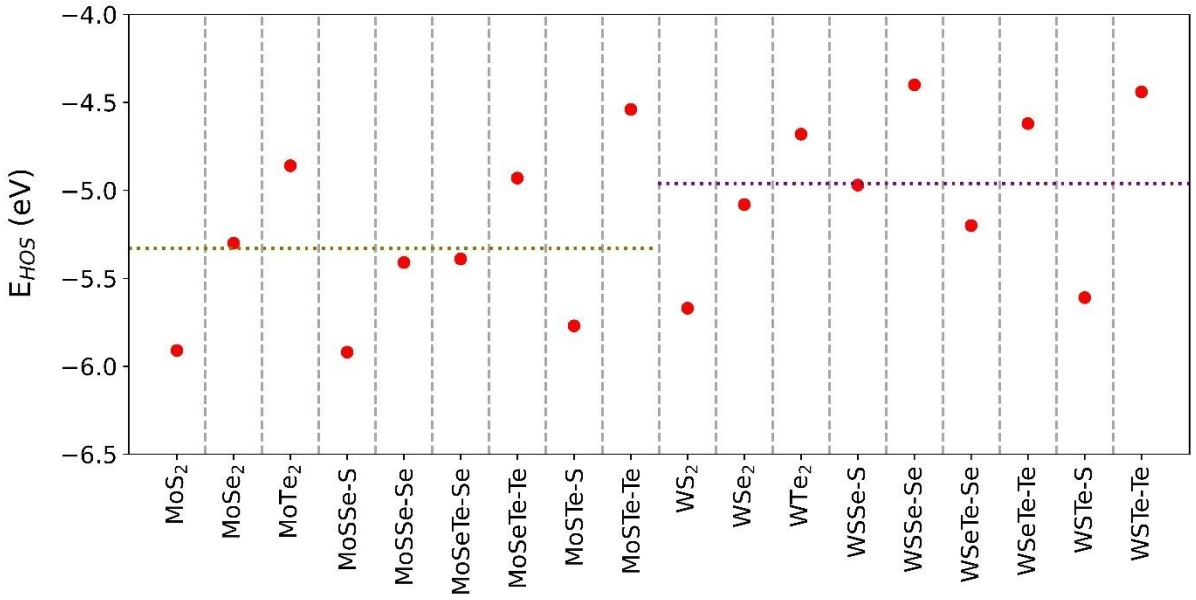
Supplementary material



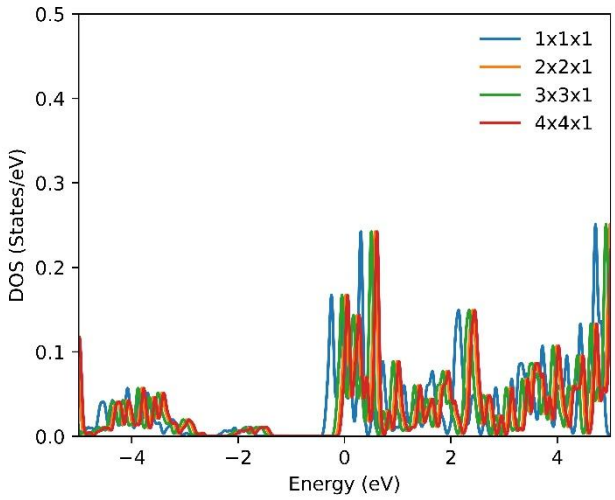
SFig. 1 The variation of the cohesive energy E_{coh} of various TMD SLs.



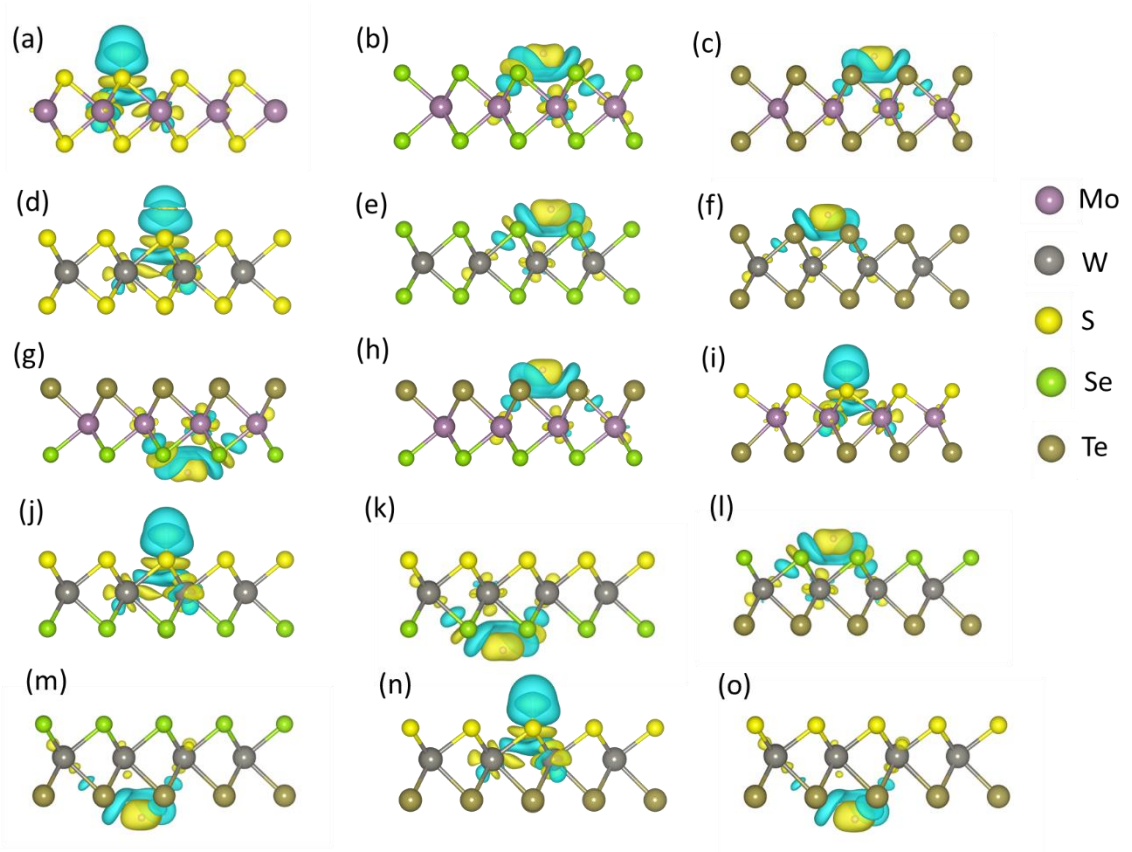
SFig. 2 The electrostatic potential of the Janus structures, MoSSe, MoSeTe, MoSTe, WSSe, WSeTe and WSTe. The red-colored dashed lines indicate the vacuum levels corresponding to the two sides of the Janus SLs.



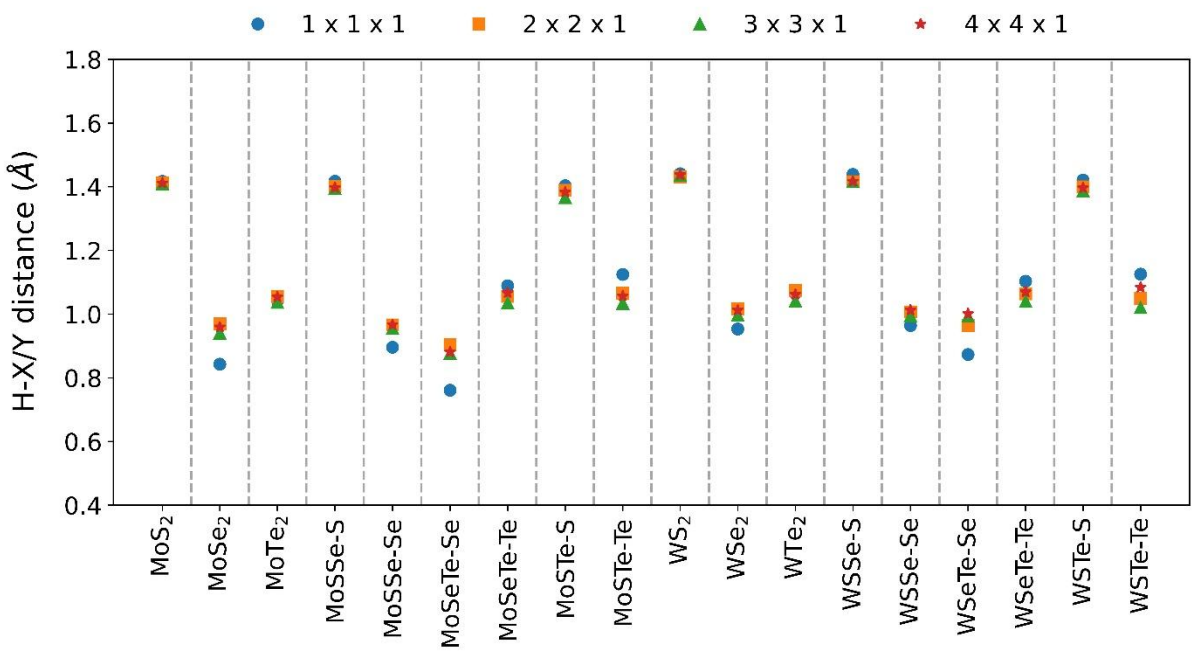
SFig. 3 The variation of the highest occupied state, E_{HOS} of various TMD SLs. The olive and purple-colored dotted lines indicate the average E_{HOS} values of the Mo and W based TMDs.



SFig. 4: The contribution of the H atom 1s orbital density of states (DOS) in the MoS_2 SL under four distinct coverage ratios.



SFig. 5 Calculated differential charge densities when the H is adsorbed on (a) MoS_2, (b) $MoSe_2$, (c) $MoTe_2$ (d) WS_2, (e) WSe_2, (f) WTe_2, (g) MoSeTe-Se, (h) MoSeTe-Te, (i) MoSTe-S, (j) WSSe-S, (k) WSSe-Se, (l) WSeTe-Se, (m) WSeTe-Te, (n) WSTe-S and (o) WSTe-S surfaces respectively. The yellow and cyan-colored isosurfaces around the atoms represent the charge accumulation and depletion respectively. The isosurface is projected at 0.002 e/Å^3.



SFig. 6 The vertical H-X/Y distances between the adsorbed H atom and the surface of adsorption in all the considered regular and Janus TMDs. The energetically favorable sites are chosen for the study i.e., C-site for S surfaces and M-site for Se and Te surfaces.

STable. 1 The calculated adsorption energy E_{ads}, charge on H atom and the average H-X/Y vertical distances of the considered TMDs.

System	Supercell	Adsorption energy (eV)	Charge on H (e)	H-X/Y distance (Å)
MoS_2	1×1×1	1.6640	-0.018	1.4167
	2×2×1	1.6600	0.020	1.4126
	3×3×1	1.8310	0.041	1.4094
	4×4×1	1.8300	0.031	1.4112
$MoSe_2$	1×1×1	1.7160	-0.010	0.8426
	2×2×1	1.7407	-0.020	0.9691
	3×3×1	1.8770	-0.030	0.9400
	4×4×1	1.9470	-0.030	0.9592
$MoTe_2$	1×1×1	1.4920	-0.200	1.0421
	2×2×1	1.5700	-0.210	1.0556
	3×3×1	1.6820	-0.220	1.0370
	4×4×1	1.7210	-0.220	1.0538
MoSSe-S	1×1×1	1.5430	-0.016	1.4172
	2×2×1	1.4500	0.044	1.4020
	3×3×1	1.6060	0.054	1.3946
	4×4×1	1.6190	0.040	1.3970
MoSSe-Se	1×1×1	1.6840	-0.010	0.8959
	2×2×1	1.7740	-0.020	0.9658

	3×3×1	1.8810	-0.040	0.9550
	4×4×1	1.9480	-0.030	0.9669
MoSeTe-Se	1×1×1	1.7570	-0.009	0.7607
	2×2×1	1.6760	-0.023	0.9035
	3×3×1	1.8310	-0.022	0.8771
	4×4×1	1.7700	-0.022	0.8807
MoSeTe-Te	1×1×1	1.4620	-0.217	1.0893
	2×2×1	1.6390	-0.225	1.0561
	3×3×1	1.6920	-0.239	1.0349
	4×4×1	1.7307	-0.236	1.0668
MoSTe-S	1×1×1	1.2259	-0.039	1.4031
	2×2×1	1.0850	0.040	1.3888
	3×3×1	1.1460	0.048	1.3661
	4×4×1	1.2100	0.041	1.3825
MoSTe-Te	1×1×1	1.4505	-0.225	1.1244
	2×2×1	1.6780	-0.237	1.0654
	3×3×1	1.6960	-0.247	1.0336
	4×4×1	1.7270	-0.245	1.0567
WS_2	1×1×1	1.9430	-0.014	1.4408
	2×2×1	1.8920	0.019	1.4314
	3×3×1	2.1070	0.042	1.4361
	4×4×1	2.1020	0.028	1.4367
WSe_2	1×1×1	1.8650	-0.047	0.9532
	2×2×1	1.9250	-0.055	1.0164
	3×3×1	2.0340	-0.064	0.9974
	4×4×1	2.1230	-0.064	1.0118
WTe_2	1×1×1	1.5700	-0.216	1.0647
	2×2×1	1.6850	-0.229	1.0752
	3×3×1	1.7720	-0.240	1.0410
	4×4×1	1.8280	-0.239	1.0614
WSSe-S	1×1×1	1.7890	-0.036	1.4386
	2×2×1	1.6740	0.027	1.4161
	3×3×1	1.8940	0.033	1.4160
	4×4×1	1.9020	0.023	1.4167
WSSe-Se	1×1×1	1.8150	-0.047	0.9641
	2×2×1	1.9550	-0.052	1.0070
	3×3×1	2.0300	-0.068	0.9944
	4×4×1	2.1160	-0.066	1.0123
WSeTe-Se	1×1×1	1.9340	-0.037	0.8730
	2×2×1	1.8690	-0.053	0.9645
	3×3×1	2.0160	-0.055	0.9934
	4×4×1	2.0970	-0.057	1.0014

WSeTe-Te	1×1×1	1.5290	-0.225	1.1032
	2×2×1	1.7500	-0.238	1.0641
	3×3×1	1.7650	-0.251	1.0402
	4×4×1	1.8250	-0.247	1.0689
WSTe-S	1×1×1	1.4970	-0.041	1.4209
	2×2×1	1.3130	0.015	1.3995
	3×3×1	1.5040	-1.182	1.3872
	4×4×1	1.5210	-1.182	1.3964
WSTe-Te	1×1×1	1.5130	-0.233	1.1253
	2×2×1	1.7800	-0.259	1.0484
	3×3×1	1.7670	-0.255	1.0205
	4×4×1	1.8150	-0.254	1.0840

STable. 2 The two descriptors $E_{TMD\text{-}H}$ and $E_{H\text{-}H}$ used in training the ML models for all the considered TMD SLs.

System	Supercell	$E_{TMD\text{-}H}$ (eV)	$E_{H\text{-}H}$ (eV)
MoS_2	1×1×1	-0.0190	-14.8826
	2×2×1	-0.0190	-3.5935
	3×3×1	-0.0190	-2.2443
	4×4×1	-0.0190	-1.9726
$MoSe_2$	1×1×1	0.0030	-11.7365
	2×2×1	0.0030	-2.9987
	3×3×1	0.0030	-2.0399
	4×4×1	0.0030	-1.8746
$MoTe_2$	1×1×1	0.0011	-15.7526
	2×2×1	0.0011	-3.9386
	3×3×1	0.0011	-2.4368
	4×4×1	0.0011	-1.7766
MoSSe-S	1×1×1	1.8575	-15.5620
	2×2×1	1.8575	-3.7731
	3×3×1	1.8575	-2.3259
	4×4×1	1.8575	-2.0171
MoSSe-Se	1×1×1	-1.8527	-15.5620
	2×2×1	-1.8527	-3.7730
	3×3×1	-1.8527	-2.3258
	4×4×1	-1.8527	-2.0171
MoSeTe-Se	1×1×1	3.2163	-10.0990
	2×2×1	3.2163	-2.7026
	3×3×1	3.2163	-1.9479
	4×4×1	3.2163	-1.8493
MoSeTe-Te	1×1×1	-4.8343	-10.0992

	2×2×1	-4.8343	-2.7026
	3×3×1	-4.8343	-1.9479
	4×4×1	-4.8343	-1.8494
MoSTe-S	1×1×1	5.0388	-15.1140
	2×2×1	5.0388	-3.7254
	3×3×1	5.0388	-2.3260
	4×4×1	5.0388	-1.8494
MoSTe-Te	1×1×1	-5.0358	-15.1150
	2×2×1	-5.0358	-3.7254
	3×3×1	-5.0358	-2.3258
	4×4×1	-5.0358	-0.0142
WS_2	1×1×1	0.0022	-17.7429
	2×2×1	0.0022	-4.2141
	3×3×1	0.0022	-2.4879
	4×4×1	0.0022	-2.0949
WSe_2	1×1×1	0.0035	-16.6503
	2×2×1	0.0035	-4.0391
	3×3×1	0.0035	-2.4424
	4×4×1	0.0035	-2.0803
WTe_2	1×1×1	0.0013	-13.4952
	2×2×1	0.0013	-3.4581
	3×3×1	0.0013	-2.2491
	4×4×1	0.0013	-1.9960
WSSe-S	1×1×1	1.9816	-17.5806
	2×2×1	1.9816	-4.2096
	3×3×1	1.9816	-2.4978
	4×4×1	1.9816	-2.1043
WSSe-Se	1×1×1	-1.9759	-17.5806
	2×2×1	-1.9759	-4.2096
	3×3×1	-1.9759	-2.4979
	4×4×1	-1.9759	-2.1047
WSeTe-Se	1×1×1	3.6166	-15.8278
	2×2×1	3.6166	-3.9108
	3×3×1	3.6166	-2.4098
	4×4×1	3.6166	-2.0710
WSeTe-Te	1×1×1	-3.6117	-15.8278
	2×2×1	-3.6117	-3.9108
	3×3×1	-3.6117	-2.4098
	4×4×1	-3.6117	-2.0710
WSTe-S	1×1×1	5.4877	-14.4136
	2×2×1	5.4877	-3.5805
	3×3×1	5.4877	-2.2698
	4×4×1	5.4877	-1.9962
WSTe-Te	1×1×1	-5.4841	-14.4136

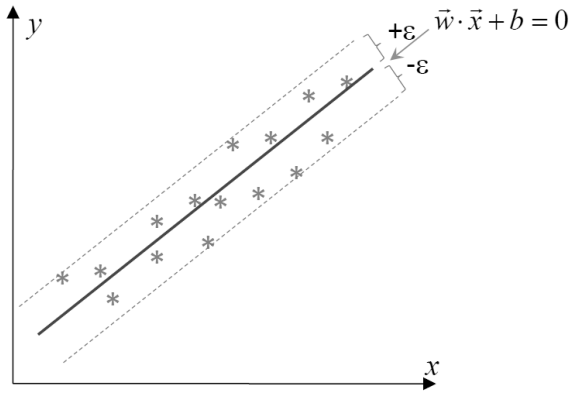
	2×2×1	-5.4841	-3.5805
	3×3×1	-5.4841	-2.2698
	4×4×1	-5.4841	-1.9962

ML approaches:

STable. 3 The comparison of the performance of the three model LR, SVR and MLPR for the individual S-, Se- and Te-surfaces with and without the inclusion of the 1×1×1 cell data.

Data	**LR**		**SVR**		**MLPR**	
	RMSE	R^2	RMSE	R^2	RMSE	R^2
S-surfaces Excluding 1×1×1	0.0386	0.988	0.0644	0.919	0.0115	0.997
Including 1×1×1	0.0900	0.803	0.0891	0.904	0.0297	0.989
Se-surfaces Excluding 1×1×1	0.0423	0.832	0.0578	0.687	0.0251	0.941
Including 1×1×1	0.0846	-0.225	0.0826	-0.168	0.0329	0.938
Te-surfaces Excluding 1×1×1	0.0335	0.419	0.0512	-1.075	0.0182	0.739
Including 1×1×1	0.0235	0.951	0.0656	0.252	0.0414	0.707

Support vector regression (SVR):



SFig. 7 Illustration of SVR in 2D-plane. Here, $\bar{x}$ denotes a point, (x, y) and $\overline{w} = (w_1, w_2)$ denotes the corresponding coefficients. Regression line: $\overline{w} \cdot \bar{x} + b = w_1 x + w_2 y + b = 0$.

In this figure, a regression line ($\overline{w} \cdot \bar{x} + b = 0$.) and the maximum deviation ($\pm\epsilon$) lines above and below the regression line (that forms a tube in 2D-plane) are shown. In this example, $\bar{x} = (x, y)$ and $\overline{w} = (w_1, w_2)$ denote a 2D-point and its coefficients, respectively.

The training points which are on the boundary of the tube are called support vectors. The final regression line is represented using the support vectors as

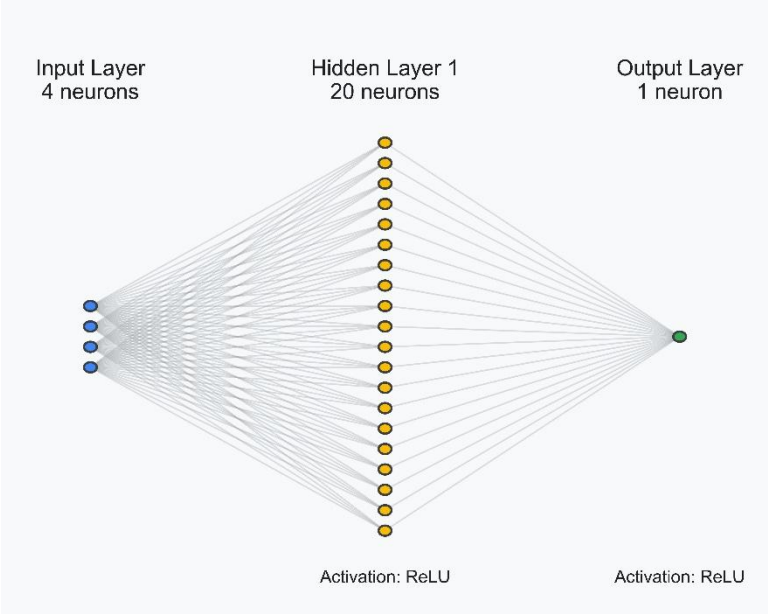
$$f(x) = \sum_{i=1}^{N} \alpha_i (x_i, x) + b$$

Here, x_i are support vectors and α_i are the corresponding multiplier coefficients. If the input data is not linearly separable, the data points are projected or transformed to a higher dimensional space using a kernel function, ϕ, to make them linearly separable in the transformed space. The regression function in this case is given by:

$$f(x) = \sum_{i=1}^{N} \alpha_n (\phi(x_i) \,.\, \phi(x)\,) + b$$

The dot product in the above equation can be interpreted as a kernel function, $K(x_i, x) = \phi(x_i) \cdot \phi(x)$. The kernel functions can be in a linear or non-linear form. Standard non-linear kernels are polynomial and radial basis functions (RBF).

Multilayer perceptron regression (MLPR):

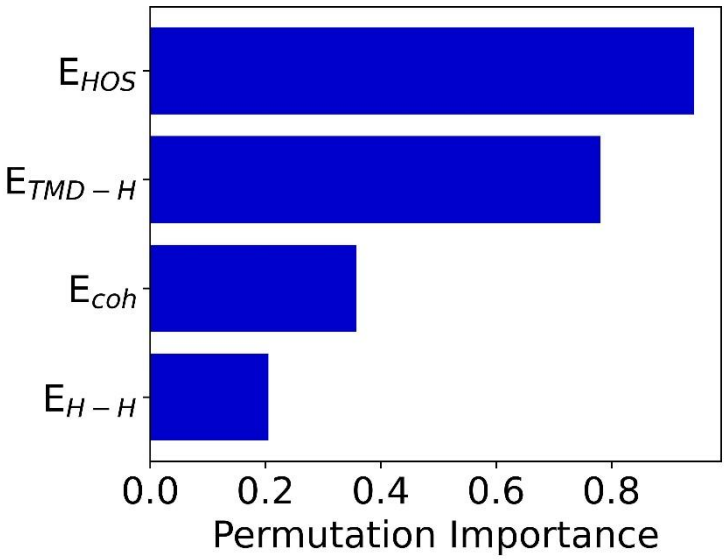


SFig. 8: The schematic of the feedforward neural network with an input layer, one hidden layer, and an output layer.

In the present model a four-dimensional input (descriptors) is fed to the network. The hidden layer has twenty neurons and the output layer has a single neuron for the output variable (H adsorption energy) to be learnt. To incorporate non-linearity into the model, we use the Rectified Linear Unit (ReLU) as the activation function at both the input to the hidden layer and the output layer.

STable 4: The performance metrics R^2 score and RMSE for the MLPR model with the variation of the number of neurons in the hidden layer.

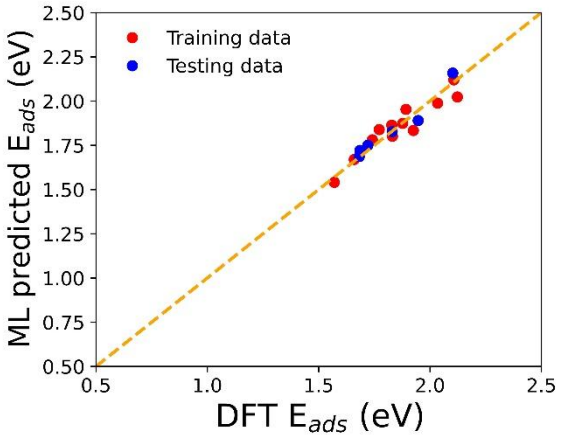
No of neurons in hidden layer	**Training data R^2**	**Testing data R^2**	**Training data RMSE (eV)**	**Testing data RMSE (eV)**
4	0.78	0.83	0.0926	0.1104
8	0.89	0.87	0.0648	0.0950
10	0.96	0.97	0.0374	0.0461
12	0.96	0.98	0.0372	0.0406
20	0.97	0.98	0.0360	0.0277
40	0.97	0.99	0.0349	0.0308
60	0.97	0.98	0.0336	0.0361
80	0.95	0.96	0.0423	0.0503
100	0.97	0.98	0.0336	0.0338



SFig. 9 The importance of each descriptor in predicting the H atom adsorption energy E_{ads} on the TMD surfaces with the MLPR model.

STable 5: The performance metrics RMSE and R2 score for the MLPR model when the 36, 72 and 144 synthetic data points are added to the original data.

Data	**Training data**		**Testing data**	
	RMSE (eV)	R^2	RMSE (eV)	R^2
72 + 36	0.0484	0.958	0.0517	0.902
72 + 72	0.0456	0.958	0.0490	0.920
72 + 144	0.0493	0.937	0.0482	0.954



SFig. 10. The DFT calculated verses ML predicted E_{ads} with the linear regression model by considering only the regular TMDs.