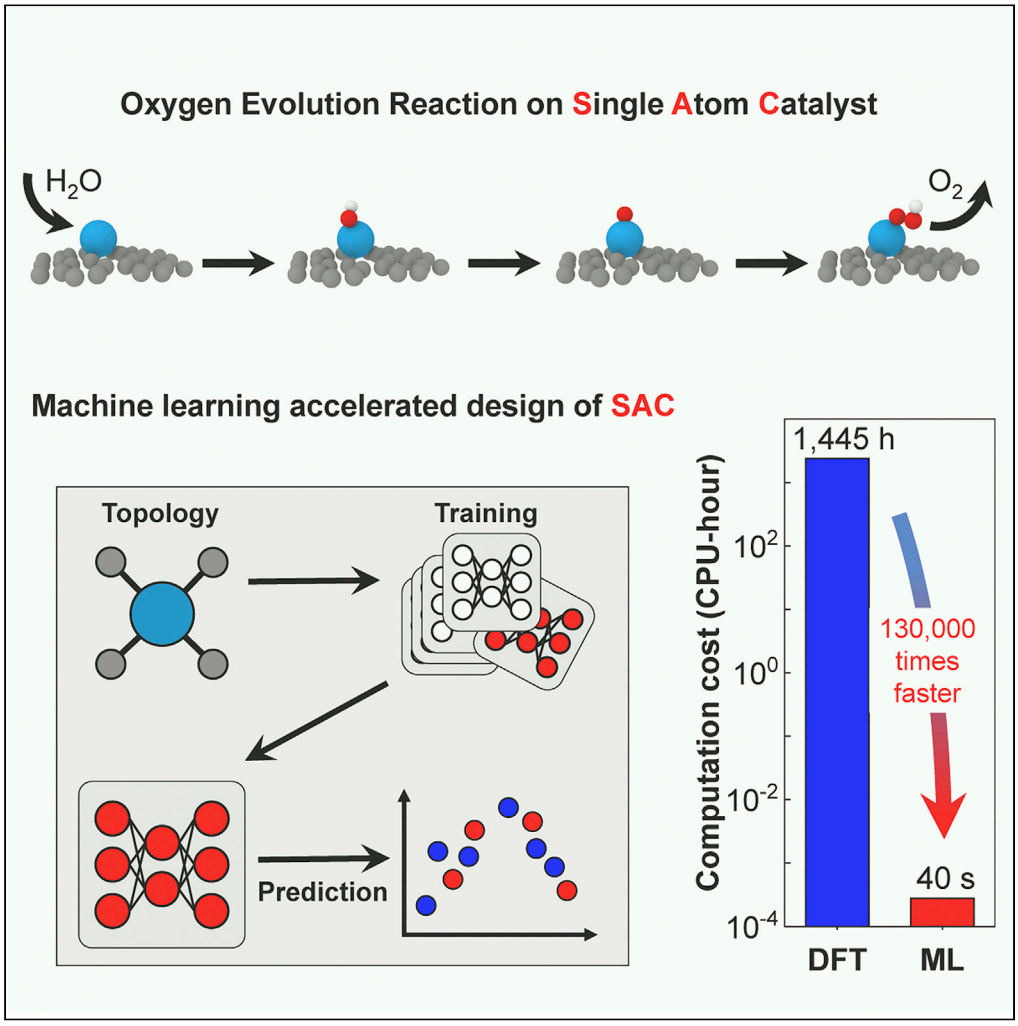
iScience

CellPress
OPEN ACCESS

Article

Machine learning-accelerated prediction of overpotential of oxygen evolution reaction of single-atom catalysts



Lianping Wu, Tian Guo, Teng Li

lit@umd.edu

Highlights

We present a topology-based machine learning (ML) approach to predict OER activity

The prediction by the ML model is of high precision (relative error of 6.49%).

The ML model is 130,000 times faster than pure density function theory calculation

Wu et al., iScience 24, 102398
May 21, 2021
https://doi.org/10.1016/j.isci.2021.102398

iScience

CellPress
OPEN ACCESS

Article

Machine learning-accelerated prediction of overpotential of oxygen evolution reaction of single-atom catalysts

Lianping Wu,[1,2] Tian Guo,[1,2] and Teng Li[1,3,*]

SUMMARY

The oxygen evolution reaction (OER) is a critical reaction for energy-related applications, yet suffers from its slow kinetics and large overpotential. It is desirable to develop effective OER electrocatalysts, such as single-atom catalysts (SACs). Here, we demonstrate machine learning (ML)-accelerated prediction of OER overpotential of all transition metals. Based on density functional theory (DFT) calculations of 15 species of SACs, we design a topological information-based ML model to map the OER overpotentials with atomic properties of the corresponding SACs. The trained ML model not only yields remarkable prediction precision (relative error of 6.49%) but also enables a 130,000-fold reduction of prediction time in comparison with pure DFT calculation. Furthermore, an intrinsic descriptor that correlates the overpotential of an SAC with its atomic properties is revealed. The approach and results from this study can be readily applicable to screen other SACs and significantly accelerate the design of high-performance catalysts for many other reactions.

INTRODUCTION

The economical, highly efficient, eco-friendly energy storage and conversion systems such as water electrolyzers and metal-air batteries (Luo et al., 2014) have attracted extensive interests due to their merits of clean and sustainable ways in consuming molecular hydrogen (H_2). In these electrochemical reaction-governed systems, the half-reaction known as the oxygen evolution reaction (OER) plays a substantial role (Amiinu et al., 2018; Li and Lu, 2017). The commonly used electrocatalysts for OER are platinum and noble metal oxides (Chen et al., 2014; Wang et al., 2015). These precious-metal-group materials are of high cost and scarcity and thus not suitable for large-scale (Bai et al., 2016) and sustainable applications (Greeley et al., 2009). Hence, there exist significant efforts to develop cheaper catalysts including metal-free catalysts and alloys with cheaper elements. In addition, downsizing the precious-metal-group materials to provide the highest number of active sites in the catalyst and achieve the maximum utilization efficiency is also widely studied (Qiao et al., 2011).

Recently, transition metal single-atom catalysts (SACs) dispersed on a carbon substrate have attracted extensive research attention as a new frontier in catalysis science (Li et al., 2021; Liu et al., 2018; Liu and Corma, 2018; Sun et al., 2019; Yan et al., 2018). Transition metals have an incomplete *d*-orbital that allows the metal to facilitate exchange of electrons, thus can both give and accept electrons easily. By taking advantage of the metal *d* orbitals, the transition metal atoms bonded to the defective sites of a graphene can accelerate reactions by means of coordination, ligand exchange, insertion, and elimination, leading to the cleavage or formation of OH, O, and OOH bonds (Jiang et al., 2017; Xu et al., 2018). Transition metal SACs possess unique physical and chemical properties that are distinct from conventional nanoparticles (NPs) and metal catalysts, including high selectivity (Kwon et al., 2017; Lee et al., 2020; Li et al., 2019), tunable high activity (Wang et al., 2019), and maximum atomic efficiency (Zhang et al., 2019). Graphene has received enormous interest in the field of electrochemistry as a promising catalytic support due to its lightweight, low cost, adjustable porosity, high chemical and thermal stability, and controllable chemical properties by heteroatom doping (Wang et al., 2012, 2013, 2014a, 2014b). Furthermore, the surface of carbon nanomaterials (e.g., NPs and nanowires) shows high capacity for stabilizing metal SACs, resulting from the existence of defects in the surface (Wei et al., 2018; Wu et al., 2020a, 2020b).

[1]Department of Mechanical Engineering, University of Maryland, College Park, MD, USA 20742

[2]These authors contributed equally

[3]Lead contact

*Correspondence: lit@umd.edu

https://doi.org/10.1016/j.isci.2021.102398

Promise of transition metal SACs on a carbon surface aside, it remains a grand challenge to search for highly efficient transition metal SACs using trial and error approaches in a reasonable timescale from hundreds of possible combinations of transition metal species and various defect structures on the carbon surface.

Machine learning (ML) algorithms with multiple processing layers to enable data learning via multiple levels of abstraction have begun to be utilized in materials science research, e.g., identifying structural flow defects in disordered solids (Cubuk et al., 2015), modeling and designing composite materials (Chen and Gu, 2019; Gu et al., 2018a, 2018b), discovering inorganic-organic hybrid materials (Raccuglia et al., 2016), and predicting the new stable structure of quaternary Heusler compounds (Kim et al., 2018). In searching for high performance catalysts, ML has been used to establish the correlations of physical properties and adsorption strength of the reaction intermediates (O'Connor et al., 2018) and to identify the relationships between the intermediate adsorption strengths and the performance of the catalyst (Ma et al., 2015) (Lin et al., 2020). Recently, the ML algorithm is also used to depict the underlying pattern of the physical properties of 104 graphene-supported SACs and their limiting potentials toward the oxygen reduction reaction/OER/hydrogen evolution reaction and predict the catalytic performance of 260 other graphene-supported metal-nitrogen/carbon systems (Lin et al., 2020). However, the amount of training data of the ML model in Ref. 40 is large and thus requires significant computation cost.

Aiming to address the above challenges, here we report a simulation-based, ML-accelerated prediction of the OER overpotential of SACs of all non-radioactive transition metals on a graphene. We first perform density functional theory (DFT) calculations to evaluate the overpotential of OER of 15 common transition metals (i.e., Ti, V, Cr Mn, Fe, Co, Ni, Cu, Mo, Ru, Pd, Ag, Pt, Au, and Zn) stabilized at two types of vacancy defects on graphene. Then, we use the DFT results from the 15 transition metal species as the training and testing data sets for a topological information-based ML algorithm, which in turn can predict the OER catalytic performance of SACs of all non-radioactive transition metals. We define an intrinsic descriptor to characterize the OER catalytic activities of various SACs on graphene. The ML-accelerated design is shown to drastically reduce the computation time to predict and evaluate the OER catalytic performance of SACs of all non-radioactive transition metals, 130,000 times faster than the approach based on pure DFT calculation, yet with an encouraging prediction precision (<6.49% error). The simulation-based, ML-accelerated prediction strategy offers quantitative guidance for rational selection of transition metals to fabricate SACs with desirable electrocatalyst activity and at a significantly reduced cost.

RESULTS

OER performance of transition metal SACs

Previous experimental studies have shown that the realistic active sites of SACs supported on graphene are actually isolated metal atoms coordinated by carbon atoms (Yao et al., 2019). Based on our previous study (Wu et al., 2020a), we construct two models, in which a metal SAC is doped on a graphene with different coordination environments: the SAC at a single vacancy site (SV site) with three carbon neighbors and the SAC at a double vacancy site (DV site) with four carbon neighbors, respectively. The detailed geometric and electronic structures of the metal atoms on graphene are shown in Figures S1–S16. According to the typical experimental electrocatalytic conditions for SACs supported on a graphene, we evaluate the OER performance of the SACs in an acidic electrolyte (Calle-Vallejo et al., 2015) (see transparent methods for details). Consistent with other studies (Liao et al., 2012), we assume a four-step OER mechanism that proceeds through OOH*, O*, and OH* (the asterisk denotes the adsorption site). The reaction scheme with the intermediates in the OER process on SACs at the SV site and DV site is shown in Figure 1A.

The calculated OER energy diagrams at 1.23 V for the SACs of 14 transition metals at the SV site and the DV site following the suggested reaction pathways are presented in Figures 1B–1E. We choose to show the OER energy diagrams at 1.23 V because the reaction free energies of all four proton transfer steps should be the same at zero potential (4.92 V/4 = 1.23 V) for an ideal catalyst. From the OER energy diagram, we can directly evaluate the overpotential η^{OER} using the limiting reaction barrier determined from the free energy of the rate-determining step (RDS). For nine species of SACs (Co, Cr, Cu, Fe, Mn, Mo, Ti, V, and Ru) on the SV site, the RDS is the last step to form the O_2 with a limiting reaction barrier ranging from 0.52 to 1.43 eV. For Ag and Zn SACs on the SV site, the RDS is the oxidation of OH* to O* with a limiting reaction barrier of 0.72 eV and 1.18eV, respectively. For other species of SACs (Au, Ni, Pd, and Pt) on the SV site, the RDS is the oxidation of O* to OOH* with limiting barrier ranging from 0.99 to 1.06 eV. For the species of SAs on the DV site, the RDS is the oxidation of O* to OOH* (Co, Cr, Fe, Mn, Mo, Pd, Pt, Ru, and V), or is the oxidation of

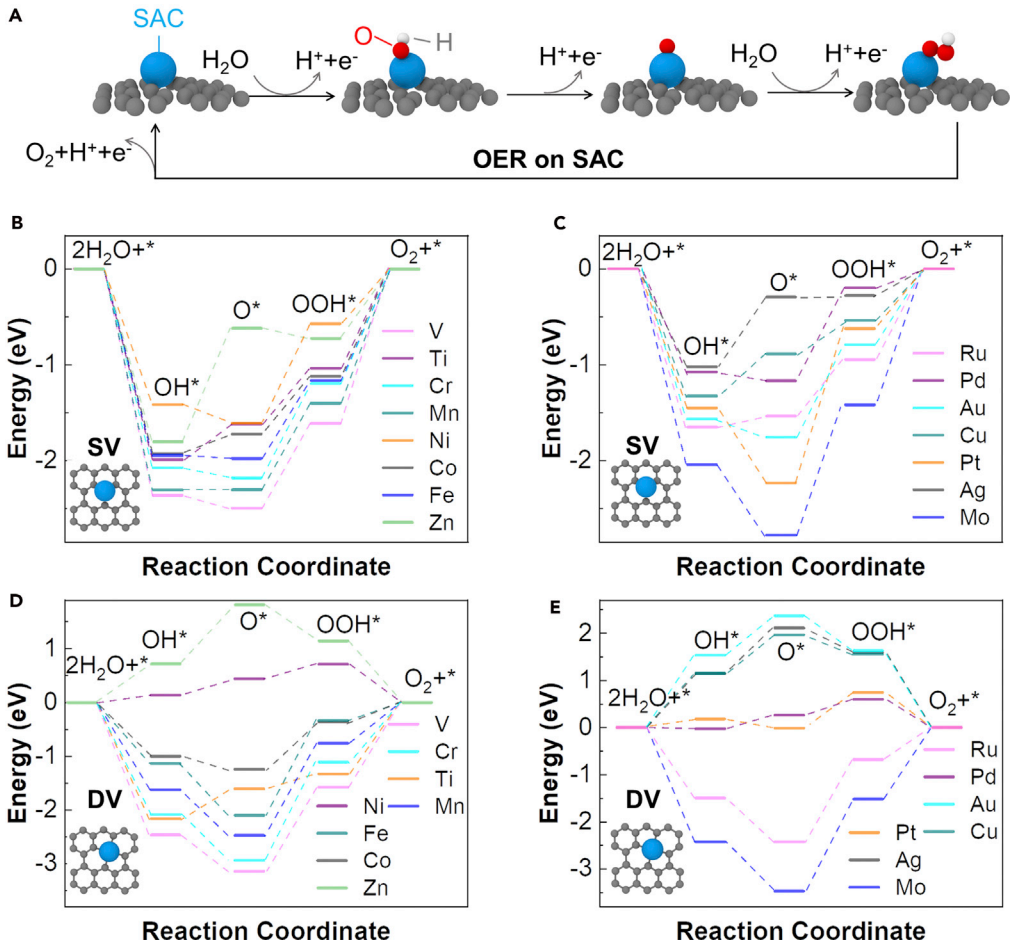


Figure 1. Evaluation of catalytic activity of transition metal SACs

(A) The reaction scheme with the intermediates in the OER process of an SAC on an SV site.

(B–E) Free energy diagram at 1.23 V for OER over the SV site (B, C) and the DV site (D, E).

OH* to O* (Ni and Zn), or is the first step to form OH* (Ag and Au), or is the last step to form O_2 (Ti). The limiting reaction barrier for the SAs on DV sites range from 0.29 (Ni) to 1.94 eV (Mo).

The DFT calculated overpotential η^{OER} of transition metal SACs as shown in Figure 1 offers quantitative guidance on the catalytic performance of SACs of various species in the OER process. However, such a calculation requires tremendous computation cost. For example, on average, it takes ~1,445 CPU-hour (~36.125 CPU-hour in parallel on a 40-core supercomputer [Intel E5-2680, 2.80GHz]) to calculate one overpotential η^{OER} value using the Vienna Ab Initio Simulation Package (VASP). The high computation cost poses a significant challenge to a comprehensive study of the OER activity of all transition metals, a highly desirable parameter to close the loop of the rational selection and design of high-performance SACs.

Topological information-based ML algorithm

To address the challenge of prohibitive computing expense to screen all transition metal SACs and evaluate their OER performance using pure DFT calculations, we introduce a topological information-based ML algorithm to predict the OER overpotential of all transition metal SACs. Figure 2 illustrates the implementation of the ML-based prediction, which contains three steps:Step 1

Data generation. First, the OER overpotential of 15 transition metal SACs obtained by DFT calculations is collected to serve as the target data. Since the electronic characteristics of catalytic centers influence the electron transfer and reaction energy in OER (Nayak et al., 2019; Qiu et al., 2015; Wang et al., 2017), five parameters of transition metals in two categories (Table S1), including atomic dimension and structure (atomic mass *M*,

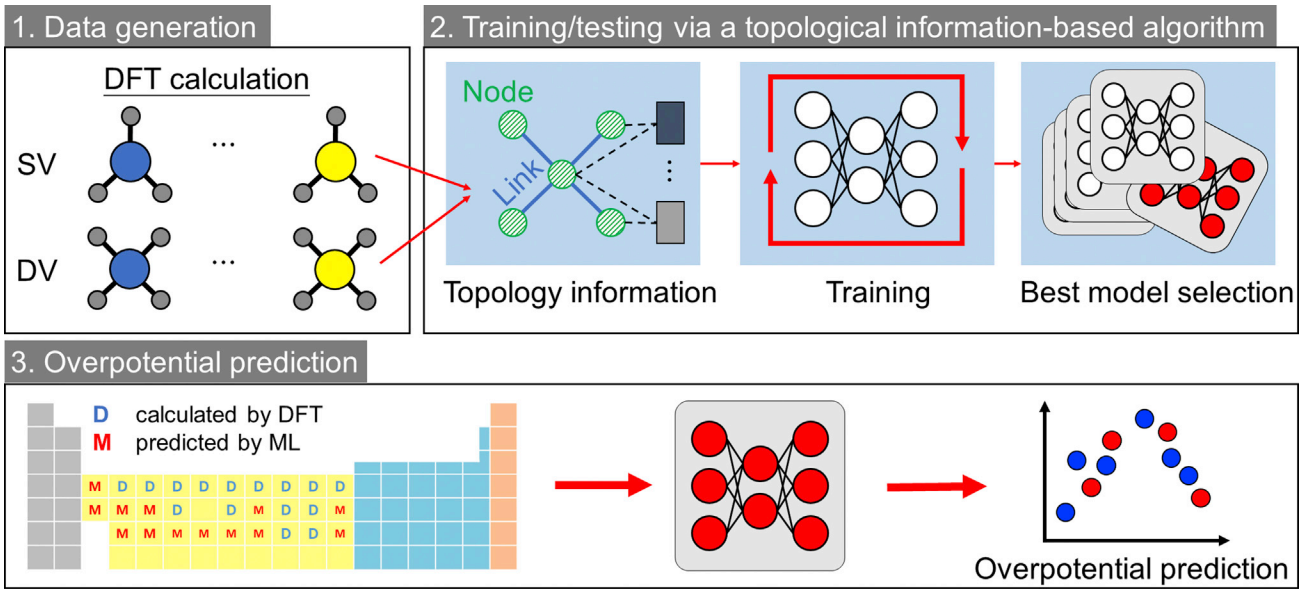


Figure 2. Diagram describing the training strategy for ML models involving (1) data generation, (2) training/test via a topological information-based algorithm, and (3) overpotential prediction workflow.

atomic radius r_M, and *d*-electron count θ_d), and reactivities (electronegativity E_M, electron affinity E_A, and first ionization energy E_i), are identified as the input data. The input data ($\boldsymbol{x}_i = [M, r_M, \theta_d, E_M, E_A]$) and target data ($\eta^{OER}$) of the 15 studied transition metals form a data set for the ML model.Step 2

Training/testing via a topological information-based algorithm. The ML model is composed of multiple processing layers to extract the features of input data with multiple levels of abstraction (Lillicrap et al., 2016). The data set containing 15 input-target pairs obtained in step 1 is split into two sub-data sets: a testing sub-data set containing four input-target pairs corresponding to Mn/Mo and Mn/Cr in SV and DV sites, respectively, and1 a training sub-data set containing the remaining 26 input-target pairs. The partition method of the data set into two sub-data sets is completely random.

In consideration of a single transition metal atom supported on graphene with different coordination environment, we use the topological information to represent the geometric spatial structure between the transition metal atom and carbon atoms. Specifically, we set the atomic parameters identified in step 1 as the node information and build the links according to the bonding states. The established topological information is identified as the inputs of the ML model, which contains six convolutional layers and three full connection layers. The convolutional layers are used to extract multiple-level features between the transition metal atom and carbon atoms. The outputs of the last convolutional layer serve as the inputs of the full connection layers, which make a regression to predict the overpotential of the transition metal atom.

We implement the ML model in PyTorch (Paszke et al., 2019) and then feed the two sub-data sets into the ML model to optimize the model parameters. The training process is performed using the Adam algorithm (Kingma and Ba, 2014) with a fixed learning rate of 0.001 for 180 epochs. In each training process, we use mean squared error as the loss function (Adamson and Smith, 2018). We run 50 training processes and select the ML model with the lowest relative deviation, $e = \sum_i^n \frac{|x_i - \bar{x}|}{\bar{x}}$, for the testing sub-data set as the best performing ML model, as shown in Figure S17. To demonstrate the ML model is not overfit, we record the relative deviation during the training process, as shown in Figure S18. The ML training and testing step is carried out for the case of an SAC on an SV site and the case of an SAC on a DV site independently. At the end of the ML training and testing step, the OER overpotentials predicted by the best performing ML model agree with those obtained from the DFT calculations within an error of 6.70% and 6.49% in training and testing data sets, respectively (detailed in Section 2.3, Figure 3).Step 3

Model prediction. The best performing ML models obtained in step 2 is then used to predict the OER catalytic activities η^{OER} of an SAC of the remaining 14 transition metal species (that is, Sc, Y, Zr, Tc [Tc is radioactive but included here. The most stable isotopes of Tc have a half-life ranging from 211,000 years to 4.21 million years.], Nb, Rh, Cd, Hf, Ta, W, Re, Os, Ir, and Hg) on an SV site and a DV site on a carbon surface, respectively.

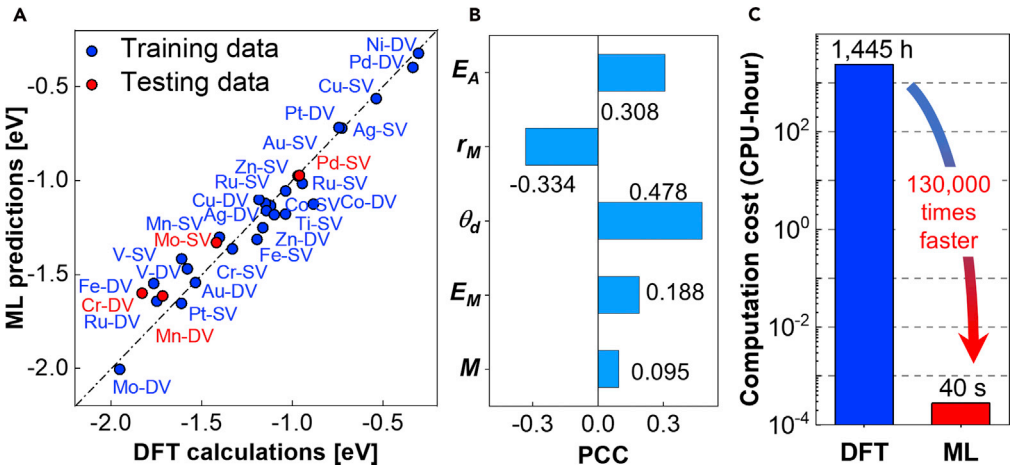


Figure 3. Machine learning drastically accelerates the prediction of OER catalytic activity

(A) Comparison between DFT and ML predicted overpotential values of η^{OER}.

(B) PCC of the five atomic properties of an SAC on an SV site and a DV site.

(C) Comparison of the average computation costs for predicting the OER catalytic activity of a transition metal SAC via pure DFT calculations and ML prediction.

ML-enabled predictions of OER catalytic activity

Figure 3A compares the ML predictions and DFT data on OER overpotential of the 12 transition metal SACs in the training set and 2 in the testing set. The values of OER overpotential of different SACs by DFT and ML are listed in Table S2. The mean relative error is 6.70% for the training data and 6.49% for the testing data, which clearly demonstrate the impressive precision of the ML predictions based on topological information-based algorithm. Further, we calculate 6 models (Rh-SV, Rh-DV, W-DV, W-SV, Hf-SV, and Hf-DV) using DFT to compare with the ML predicted value for validation. Our ML predicted errors are only 2.07%, 0.54%, 9.83%, 8.98%, 9.43%, and 12.84% for Rh-SV, Rh-DV, W-DV, W-SV, Hf-SV, and Hf-DV, respectively. To evaluate the correlation between each atomic parameter of a transition metal SAC and its OER overpotential, we calculate the Pearson correlation coefficient (PCC) to compare the importance of these five atomic parameters (Figure 3B). We find that the dominating parameters are *d*-electron count (θ_d), the atomic radius (r_M), and electron affinity (E_A), with the PCC value of 0.478, −0.334, and 0.308, respectively, followed by electronegativity (E_M) and atomic mass (M), with the PCC value of 0.188 and 0.095, respectively.

It is worth noting that the ML model can drastically accelerate the prediction of η^{OER} with a high efficiency and a high accuracy. For comparison, we time the central processing unit hours (CPU-h) to evaluate the computation cost of DFT-based calculations and ML-based predictions of η^{OER}. For DFT-based calculations, the calculation of η^{OER} for a given transition metal SAC is performed using the VASP on a 40-core supercomputer (Intel E5-2680, 2.80GHz), which on average takes ~36.125 CPU-h to calculate one η^{OER} value, equivalent to ~1,445 CPU-h on a single core CPU. By contrast, for ML-based predictions, it takes only ~40 s, starting from ML model training, to model testing, and to model predicting, on a single core CPU. In other words, ML-based predictions of the catalytic activity of transition metal SACs are more than 130,000 times faster than the DFT-based calculation (Figure 3C). Such a remarkable reduction of computation cost by ML-accelerated material discovery approach effectively overcomes the challenge of low efficiency of DFT calculations in identifying high performance transition metal SACs.

To help identify the optimal SAC, we identify an intrinsic catalytic activity descriptor to establish a "volcano" relationship between η^{OER} and the descriptor (Nørskov et al., 2009). According to the importance in predicting OER activity using the DFT calculation and ML models, we introduce an intrinsic catalytic activity descriptor as follows:

$$\phi = E_i \theta_d M \left(\frac{E_M}{r_M} + \frac{n_C E_C}{r_C} \right), \qquad \text{(Equation 1)}$$

where E_C and r_C represent the electronegativity and atomic radius of carbon, respectively. n_C represents the nearest neighbor carbon atoms.

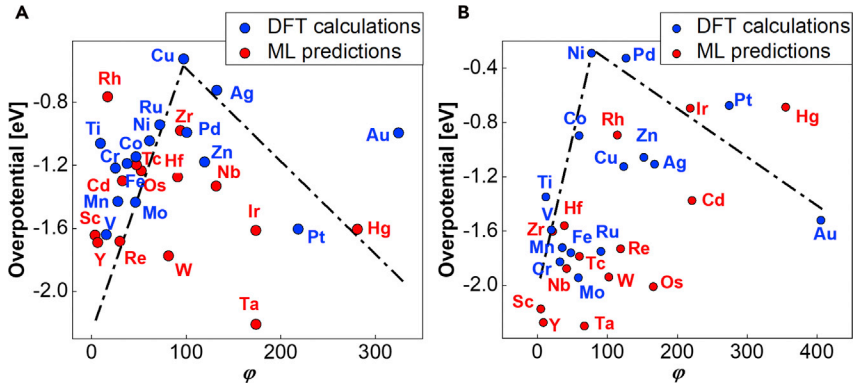


Figure 4. The volcano relationships between η^{OER} and descriptor ϕ

OER catalytic activities as a function of an intrinsic catalytic activity descriptor ϕ of SACs on an SV site (A) and a DV site (B).

Figures 4A and 4B plot the overpotential η^{OER} of the SACs of all transition metals (from both DFT calculations and ML predictions) on an SV site and a DV site on a carbon surface, respectively, as a function of ϕ, both of which show a volcano relationship. For example, for SACs on an SV site, Cu appears at the peak of the volcano (Figure 4A, i.e., with the best OER catalytic activity), while for SACs on a DV site, Rh appears at the peak of the volcano (Figure 4B). The volcano relationships between η^{OER} and ϕ shown in Figures 4A and 4B are instrumental in selecting SACs of transition metals with desirable OER catalytic activity. We further compare our calculation results with experimental measurements (Fei et al., 2018). Experiments in (Fei et al., 2018) show that the OER activity at a current density of 10 mA cm-2 (η10) of the single metal atom on doped holey graphene frameworks (NHGFs) is Ni > Co > Fe. Specifically, the Ni-NHGF catalyst shows that its potential of η10 is 331 mV, lower than that of Co-NHGF (402 mV) and Fe-NHGF (488 mV). Our DFT calculations performed on the SV site and DV site show the same trend of their OER activities (i.e., Ni > Co > Fe).

CONCLUSIONS

In summary, using a topological information-based algorithm, we built two ML models to reveal the underlying correlation of physical properties and overpotentials toward the OER of graphene-supported SACs. An intrinsic descriptor is identified to be able to well correlate with OER catalytic activity of transition metal SACs in a volcano relationship, which enables a facile selection of optimal SACs with desirable catalytic performance. The ML model featuring a topological information-based algorithm enables a 130,000-fold acceleration of the computational efficiency in calculating the OER catalytic activity of transition metal SACs. The approach and results from this study can be readily applicable to screen other SACs and significantly accelerate the design of high-performance catalysts for many other reactions.

Limitations of the study

Our work demonstrates an ML-accelerated approach for the prediction of the overpotential of OER and shows the OER overpotential of all the transition SACs on the SV site and DV site of a carbon surface. However, the influence of the substrate on the OER activity needs to be further investigated.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Teng Li (lit@umd.edu).

Materials availability

This study did not generate new reagents.

Data and code availability

This article includes all data sets/code generated or analyzed during this study.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.102398.

ACKNOWLEDGMENTS

We acknowledge the University of Maryland supercomputing resources (http://hpcc.umd.edu) and Maryland Advanced Research Computing Center (MARCC) made available for conducting the research reported in this paper.

AUTHOR CONTRIBUTIONS

T.L. conceived the research idea and supervised the research. L.W. carried out DFT. T.G. performed ML predictions. All authors contributed to result analysis and paper writing.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: November 30, 2020
Revised: February 25, 2021
Accepted: April 2, 2021
Published: May 21, 2021

REFERENCES

Adamson, A.S., and Smith, A. (2018). Machine learning and health care disparities in dermatology. JAMA Dermatol. *154*, 1247–1248.

Amiinu, I.S., Liu, X., Pu, Z., Li, W., Li, Q., Zhang, J., Tang, H., Zhang, H., and Mu, S. (2018). From 3D ZIF nanocrystals to Co–Nx/C nanorod array electrocatalysts for ORR, OER, and Zn–air batteries. Adv. Funct. Mater. *28*, 1704638.

Bai, X., Zhao, E., Li, K., Wang, Y., Jiao, M., He, F., Sun, X., Sun, H., and Wu, Z. (2016). Theoretical investigation on the reaction pathways for oxygen reduction reaction on silicon doped graphene as potential metal-free catalyst. J. Electrochem. Soc. *163*, F1496.

Calle-Vallejo, F., Tymoczko, J., Colic, V., Vu, Q.H., Pohl, M.D., Morgenstern, K., Loffreda, D., Sautet, P., Schuhmann, W., and Bandarenka, A.S. (2015). Finding optimal surface sites on heterogeneous catalysts by counting nearest neighbors. Science *350*, 185–189.

Chen, C.-T., and Gu, G.X. (2019). Machine learning for composite materials. MRS Commun. *9*, 556–566.

Chen, C., Kang, Y., Huo, Z., Zhu, Z., Huang, W., Xin, H.L., Snyder, J.D., Li, D., Herron, J.A., and Mavrikakis, M. (2014). Highly crystalline multimetallic nanoframes with three-dimensional electrocatalytic surfaces. Science *343*, 1339–1343.

Cubuk, E.D., Schoenholz, S.S., Rieser, J.M., Malone, B.D., Rottler, J., Durian, D.J., Kaxiras, E., and Liu, A.J. (2015). Identifying structural flow defects in disordered solids using machine-learning methods. Phys. Rev. Lett. *114*, 108001.

Fei, H., Dong, J., Feng, Y., Allen, C.S., Wan, C., Volosskiy, B., Li, M., Zhao, Z., Wang, Y., and Sun, H. (2018). General synthesis and definitive structural identification of MN 4 C 4 single-atom catalysts with tunable electrocatalytic activities. Nat. Catal. *1*, 63–72.

Greeley, J., Stephens, I., Bondarenko, A., Johansson, T.P., Hansen, H.A., Jaramillo, T., Rossmeisl, J., Chorkendorff, I., and Nørskov, J.K. (2009). Alloys of platinum and early transition metals as oxygen reduction electrocatalysts. Nat. Chem. *1*, 552–556.

Gu, G.X., Chen, C.-T., and Buehler, M.J. (2018a). De novo composite design based on machine learning algorithm. Extreme Mech. Lett. *18*, 19–28.

Gu, G.X., Chen, C.-T., Richmond, D.J., and Buehler, M.J. (2018b). Bioinspired hierarchical composite design using machine learning: simulation, additive manufacturing, and experiment. Mater. Horizons *5*, 939–945.

Jiang, K., Siahrostami, S., Akey, A.J., Li, Y., Lu, Z., Lattimer, J., Hu, Y., Stokes, C., Gangishetty, M., and Chen, G. (2017). Transition-metal single atoms in a graphene shell as active centers for highly efficient artificial photosynthesis. Chem *3*, 950–960.

Kim, K., Ward, L., He, J., Krishna, A., Agrawal, A., and Wolverton, C. (2018). Machine-learning-accelerated high-throughput materials screening: discovery of novel quaternary Heusler compounds. Phys. Rev. Mater. *2*, 123801.

Kingma, D.P., and Ba, J. (2014). Adam: A Method for Stochastic Optimization (arXiv), p. 14126980.

Kwon, Y., Kim, T.Y., Kwon, G., Yi, J., and Lee, H. (2017). Selective activation of methane on single-atom catalyst of rhodium dispersed on zirconia for direct conversion. J. Am. Chem. Soc. *139*, 17694–17699.

Lee, W.H., Ko, Y.-J., Kim, J.-Y., Min, B.K., Hwang, Y.J., and Oh, H.-S. (2020). Single-atom catalysts for the oxygen evolution reaction: recent developments and future perspectives. Chem. Commun. *56*, 12687–12697.

Li, J., Guan, Q., Wu, H., Liu, W., Lin, Y., Sun, Z., Ye, X., Zheng, X., Pan, H., and Zhu, J. (2019). Highly active and stable metal single-atom catalysts achieved by strong electronic metal–support interactions. J. Am. Chem. Soc. *141*, 14515–14519.

Li, Y., Hu, R., Chen, Z., Wan, X., Shang, J.-X., Wang, F.-H., and Shui, J. (2021). Effect of Zn atom in Fe-N-C catalysts for electro-catalytic reactions: theoretical considerations. Nano Res. *14*, 611–619.

Li, Y., and Lu, J. (2017). Metal–air batteries: will they be the future electrochemical energy storage device of choice? ACS Energy Lett. *2*, 1370–1377.

Liao, P., Keith, J.A., and Carter, E.A. (2012). Water oxidation on pure and doped hematite (0001) surfaces: prediction of Co and Ni as effective dopants for electrocatalysis. J. Am. Chem. Soc. *134*, 13296–13309.

Lillicrap, T.P., Cownden, D., Tweed, D.B., and Akerman, C.J. (2016). Random synaptic feedback weights support error backpropagation for deep learning. Nat. Commun. *7*, 1–10.

Lin, S., Xu, H., Wang, Y., Zeng, X.C., and Chen, Z. (2020). Directly predicting limiting potentials from easily obtainable physical properties of graphene-supported single-atom

electrocatalysts by machine learning. J. Mater. Chem. A *8*, 5663–5670.

Liu, J.-C., Tang, Y., Wang, Y.-G., Zhang, T., and Li, J. (2018). Theoretical understanding of the stability of single-atom catalysts. Natl. Sci. Rev. *5*, 638–641.

Liu, L., and Corma, A. (2018). Metal catalysts for heterogeneous catalysis: from single atoms to nanoclusters and nanoparticles. Chem. Rev. *118*, 4981–5079.

Luo, J., Im, J.-H., Mayer, M.T., Schreier, M., Nazeeruddin, M.K., Park, N.-G., Tilley, S.D., Fan, H.J., and Grätzel, M. (2014). Water photolysis at 12.3% efficiency via perovskite photovoltaics and Earth-abundant catalysts. Science *345*, 1593–1596.

Ma, X., Li, Z., Achenie, L.E., and Xin, H. (2015). Machine-learning-augmented chemisorption model for CO2 electroreduction catalyst screening. J. Phys. Chem. Lett. *6*, 3528–3533.

Nayak, S., Bhattacharjee, S., Choi, J.-H., and Lee, S.C. (2019). Machine learning and scaling laws for prediction of accurate adsorption energy. J. Phys. Chem. A *124*, 247–254.

Nørskov, J.K., Bligaard, T., Rossmeisl, J., and Christensen, C.H. (2009). Towards the computational design of solid catalysts. Nat. Chem. *1*, 37–46.

O'Connor, N.J., Jonayat, A., Janik, M.J., and Senftle, T.P. (2018). Interaction trends between single metal atoms and oxide supports identified with density functional theory and statistical learning. Nat. Catal. *1*, 531–539.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., and Antiga, L. (2019). Pytorch: An Imperative Style, High-Performance Deep Learning Library (arXiv), p. 191201703.

Qiao, B., Wang, A., Yang, X., Allard, L.F., Jiang, Z., Cui, Y., Liu, J., Li, J., and Zhang, T. (2011). Single-atom catalysis of CO oxidation using Pt 1/FeO x. Nat. Chem. *3*, 634–641.

Qiu, M., Fang, Z., Li, Y., Zhu, J., Huang, X., Ding, K., Chen, W., and Zhang, Y. (2015). First-principles investigation of the activation of CO2 molecule on TM/Cu (TM= Fe, Co and Ni) surface alloys. Appl. Surf. Sci. *353*, 902–912.

Raccuglia, P., Elbert, K.C., Adler, P.D., Falk, C., Wenny, M.B., Mollo, A., Zeller, M., Friedler, S.A., Schrier, J., and Norquist, A.J. (2016). Machine-learning-assisted materials discovery using failed experiments. Nature *533*, 73–76.

Sun, T., Xu, L., Wang, D., and Li, Y. (2019). Metal organic frameworks derived single atom catalysts for electrocatalytic energy conversion. Nano Res. *12*, 2067–2080.

Wang, G., Wang, H., Lu, X., Ling, Y., Yu, M., Zhai, T., Tong, Y., and Li, Y. (2014a). Solid-state supercapacitor based on activated carbon cloths exhibits excellent rate capability. Adv. Mater. *26*, 2676–2682.

Wang, J., Oschatz, M., Biemelt, T., Borchardt, L., Senkovska, I., Lohe, M.R., and Kaskel, S. (2012). Synthesis, characterization, and hydrogen storage capacities of hierarchical porous carbide derived carbon monolith. J. Mater. Chem. B *22*, 23893–23899.

Wang, J., Senkovska, I., Kaskel, S., and Liu, Q. (2014b). Chemically activated fungi-based porous carbons for hydrogen storage. Carbon *75*, 372–380.

Wang, J., Senkovska, I., Oschatz, M., Lohe, M.R., Borchardt, L., Heerwig, A., Liu, Q., and Kaskel, S. (2013). Highly porous nitrogen-doped polyimine-based carbons with adjustable microstructures for CO 2 capture. J. Mater. Chem. A *1*, 10951–10961.

Wang, Y.-J., Zhao, N., Fang, B., Li, H., Bi, X.T., and Wang, H. (2015). Carbon-supported Pt-based alloy electrocatalysts for the oxygen reduction reaction in polymer electrolyte membrane fuel cells: particle size, shape, and composition manipulation and their impact to activity. Chem. Rev. *115*, 3433–3467.

Wang, Y., Tang, Y.-J., and Zhou, K. (2019). Self-adjusting activity induced by intrinsic reaction intermediate in Fe–N–C single-atom catalysts. J. Am. Chem. Soc. *141*, 14115–14119.

Wang, Z., Zhao, J., and Cai, Q. (2017). CO 2 electroreduction performance of a single transition metal atom supported on porphyrin-like graphene: a computational study. Phys. Chem. Chem. Phys. *19*, 23113–23121.

Wei, S., Li, A., Liu, J.-C., Li, Z., Chen, W., Gong, Y., Zhang, Q., Cheong, W.-C., Wang, Y., and Zheng, L. (2018). Direct observation of noble metal nanoparticles transforming to thermally stable single atoms. Nat. Nanotechnol. *13*, 856–861.

Wu, L., Guo, T., and Li, T. (2020a). Rational design of transition metal single-atom electrocatalysts: a simulation-based, machine learning-accelerated study. J. Mater. Chem. A *8*, 19290–19299.

Wu, L., Hu, S., Yu, W., Shen, S., and Li, T. (2020b). Stabilizing mechanism of single-atom catalysts on a defective carbon surface. NPJ Comput. Mater. *6*, 1–8.

Xu, H., Cheng, D., Cao, D., and Zeng, X.C. (2018). A universal principle for a rational design of single-atom electrocatalysts. Nat. Catal. *1*, 339–348.

Yan, H., Su, C., He, J., and Chen, W. (2018). Single-atom catalysts and their applications in organic chemistry. J. Mater. Chem. A *6*, 8793–8814.

Yao, Y., Huang, Z., Xie, P., Wu, L., Ma, L., Li, T., Pang, Z., Jiao, M., Liang, Z., and Gao, J. (2019). High temperature shockwave stabilized single atoms. Nat. Nanotechnol. *14*, 851–857.

Zhang, J., Zhao, Y., Chen, C., Huang, Y.-C., Dong, C.-L., Chen, C.-J., Liu, R.-S., Wang, C., Yan, K., and Li, Y. (2019). Tuning the coordination environment in single-atom catalysts to achieve highly efficient oxygen reduction reactions. J. Am. Chem. Soc. *141*, 20118–20126.

iScience, Volume *24*

Supplemental information

Machine learning-accelerated prediction of overpotential of oxygen evolution reaction of single-atom catalysts

Lianping Wu, Tian Guo, and Teng Li

Supporting Information

Machine learning accelerated prediction of overpotential of oxygen evolution reaction of single atom catalysts

Lianping Wu [†], Tian Guo [†], Teng Li[*]

Department of Mechanical Engineering, University of Maryland, College Park, Maryland, USA 20742

[†]These authors contributed equally to this work

[*] Corresponding author. E-mail: lit@umd.edu

Table of Contents:

Transparent Methods

DFT calculation

All the spin-polarized density functional theory (DFT) calculations are performed using Vienna Ab initio Simulation Package (VASP) (Kresse and Hafner, 1993, 1994). The projector-augmented wave (PAW) method (Blöchl et al., 1994; Kresse and Joubert, 1999) and Perdew-Burke-Ernzerhof (PBE) (Perdew et al., 1996) functional is used. The Kohn–Sham wave functions are expanded in a plane wave basis set with a cutoff energy of 500 eV. All atoms are allowed to relax until the forces fell below 0.01 eV/Å. The Brillouin zone is sampled by the 2 × 2 × 1 Monkhorst–Pack k-point mesh. Period boundary conditions are applied in both in-plane and interlayer directions. To eliminate the interaction between periodic images of atoms, a 12.31 Å× 12.81 Å cubic supercell is used to investigate the interaction of metal atoms with graphene and a vacuum region of 15 Å is created to ensure negligible interaction between mirror images. The spin multiplicity of metal during the oxygen evolution reactions is shown in Table S3.

The oxygen evolution reactions occur via the following steps:

$$\mathrm{M} + 2\mathrm{H_2O_{(l)}} \rightarrow \mathrm{MOH} + \mathrm{H_2O} + \mathrm{H^+} + \mathrm{e^-} \quad (1)$$

$$\mathrm{MOH} + \mathrm{H_2O} + \mathrm{H^+} + \mathrm{e^-} \rightarrow \mathrm{MO} + \mathrm{H_2O} + 2\mathrm{H^+} + 2\mathrm{e^-} \quad (2)$$

$$\mathrm{MO} + \mathrm{H_2O_{(l)}} + 2\mathrm{H^+} + 2\mathrm{e^-} \rightarrow \mathrm{MOOH} + 3\mathrm{H^+} + 3\mathrm{e^-} \quad (3)$$

$$\mathrm{MOOH} + 3\mathrm{H^+} + 3\mathrm{e^-} \rightarrow \mathrm{M} + \mathrm{O_{2(g)}} + 4\mathrm{H^+} + 4\mathrm{e^-} \quad (4)$$

where M represents the preferable adsorption site for intermediates. For each step, the reaction free energy ΔG is defined as the difference between free energies of the initial and final states as calculated by the expression (Zuluaga and Stolbov, 2011)

$$\Delta G = \Delta E + \Delta \mathrm{ZPE} - T\Delta S + \Delta G_{\mathrm{U}} + \Delta G_{\mathrm{pH}}, \quad (5)$$

where ΔE is the total energy difference between reactants and products of reactions, ΔZPE is the zero-point energy correction, ΔS is the vibrational entropy change at finite temperature T, $\Delta G = -eU$, where e is the elementary charge, U is the electrode potential, and ΔG_{pH} is the correction of the H^+ free energy.

The overpotential η can be evaluated from the Gibbs free energy differences of each step as:

$$\eta = 1.23 - \max\{\Delta G_1, \Delta G_2, \Delta G_3, \Delta G_4\}, \quad (6)$$

where ΔG_1, ΔG_2, ΔG_3, and ΔG_4 are the free energy of reaction steps shown in (1) to (4).

Topological information-based algorithm

Our ML approach is designed for establishing a regression relationship between the atomic parameters and OER catalytic activity, based on the results from DFT calculations. The ML model is established to predict the OER catalytic activity of a SAC on both SV-site and DV-site. The architecture of the ML model consists of two modules: a feature extraction module and a full connection neural network (FCNN). The feature extraction module has 4 convolutional layers, 2 max-pooling layers, and 2 full connection layers. The detailed architecture is plotted in Figure S19. The features between pairs of nodes (that is, the interatomic features) are extracted via the feature extraction module, which contains 6 convolutional layers; the function of FCNN is to predict η^{OER}. The parameters in the ML model are initialized randomly with a normal distribution. Furthermore, the rectified linear unit function, $f(x) = \max(0, x)$, where x denotes the outputs of each hidden layer, is used as the activation function for each hidden layer. The two ML models are then trained using the mean square error (MSE) loss function.

Supplemental Tables (Table S1~S3)

Table S1 Atomic parameters as inputs in the topological information-based algorithm. Related to Figure 2. The first 15 elements (shaded in blue) are selected as the training/testing dataset for ML model, and η^{OER} values of the last 14 elements (shaded in orange) are predicted by the best performing ML model.

Element	Radius [pm]	*d* electron count	Electronegativity	Electron affinity [eV]	First ionization energy [eV]
Mn	197	5	1.55	0.97	7.434
Ti	187	2	1.54	-0.02	6.828
Fe	194	6	1.83	0.46	7.902
V	179	3	1.63	0.63	6.746
Pd	202	8	2.2	1.02	8.337
Co	192	7	1.88	1.06	7.881
Cr	189	5	1.66	0.97	6.767
Mo	209	5	2.16	1.18	7.092
Ru	207	7	2.2	1.51	7.361
Ni	163	8	1.91	1.62	7.64
Cu	140	10	1.9	1.8	7.726
Ag	172	10	1.93	2	7.576
Pt	209	9	2.28	2.56	9
Au	166	10	2.54	2.8	9.226
Zn	139	10	1.65	0.09	9.394
W	210	4	2.36	1.23	7.98
Zr	186	2	1.33	0.45	6.634
Nb	207	4	1.6	1.13	6.759
Rh	195	8	2.28	1.68	7.459
Cd	158	10	1.69	0.27	8.994
Hf	212	2	1.3	0.63	6.825
Ta	217	3	1.5	0.15	7.89
Re	217	5	1.9	0.38	7.88
Os	216	6	2.2	1.44	8.7
Ir	202	7	2.2	1.97	9.1
Hg	209	10	2	-0.19	10.438
Tc	209	5	1.9	0.99	7.28
Sc	211	1	1.36	-0.73	6.561
Y	219	1	1.22	-0.4	6.217
C	170	2 (P electron)	2.55	6	11.560

Table S2 DFT calculated versus ML predicted overpotential of OERs on different SACs. Related to Figure 4.

Elements	X@SV (DFT, eV)	X@SV (ML, eV)	X@DV (DFT, eV)	X@DV (ML, eV)
Mn	-1.402	-1.302	-1.716	-1.614
Ti	-1.038	-1.179	-1.331	-1.364
Fe	-1.163	-1.251	-1.765	-1.548
V	-1.611	-1.416	-1.579	-1.469
Pd	-0.961	-0.972	-0.335	-0.397
Co	-1.12	-1.134	-0.884	-1.125
Cr	-1.195	-1.313	-1.829	-1.599
Mo	-1.419	-1.330	-1.953	-2.005
Ru	-0.945	-1.016	-1.748	-1.642
Ni	-1.037	-1.055	-0.304	-0.322
Cu	-0.537	-0.5634	-1.149	-1.121
Ag	-0.727	-0.721	-1.143	-1.161
Pt	-1.612	-1.654	-0.743	-0.718
Au	-0.969	-0.975	-1.535	-1.543
Zn	-1.184	-1.101	-1.1	-1.182

Table S3. The spin multiplicity of metal on SV site and DV site in the OER reaction. Related to Figure 1. (slab: graphene with single atom, O-slab: O atom adsorbed on single atom, OH-slab: OH^- adsorbed on single atom, and OOH-slab: OOH^- adsorbed on single atom)

element	1-V site				2-V site			
	slab	**O-slab**	**OH-slab**	**OOH-slab**	**slab**	**O-slab**	**OH-slab**	**OOH-slab**
Mn	2.553	1.854	2.075	2.047	3.088	1.854	2.074	1.645
Ti	0	0	0	0	0	0	0	0
Fe	0	1.032	0.772	0	2.88	1.032	2.226	2.127
V	1	0	0	0	0.871	0	0	0
Pd	0	0	0.105	0.075	0	0	0.261	0.171
Co	0.32	0	0	0	0.909	0	0.069	0.155
Cr	2.201	1.077	1.797	1.885	2.276	1.077	1.411	1.285
Mo	1.617	0	1.033	0.814	0.652	0	0.957	0.871
Ru	0	0.574	0	0	0	0.574	0	0
Ni	0	0	0.165	0.104	0	0	0.165	0.111
Cu	0.225	0.083	0	0	0.012	0.083	0.18	0.095
Ag	0.087	0.064	0	0	0.009	0.064	0	0.024
Pt	0	0	0.135	0.123	0	0	0.096	0.1
Au	0.245	0.072	0	0	0	0.072	0.123	-0.002
Zn	0	0.002	0.013	0.014	0	0.002	0.034	-0.003

Supplemental Figures (Figure S1~S19)

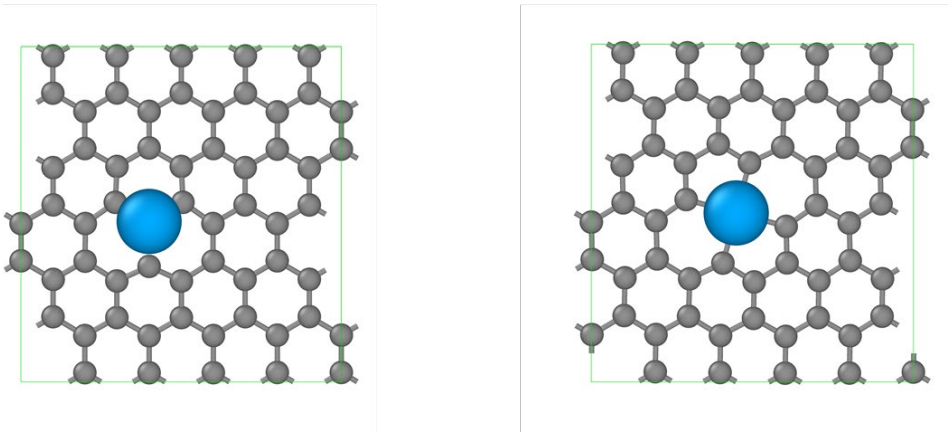


Fig. S1 Geometric structures of the metal atoms on graphene. Related to Figure 1.

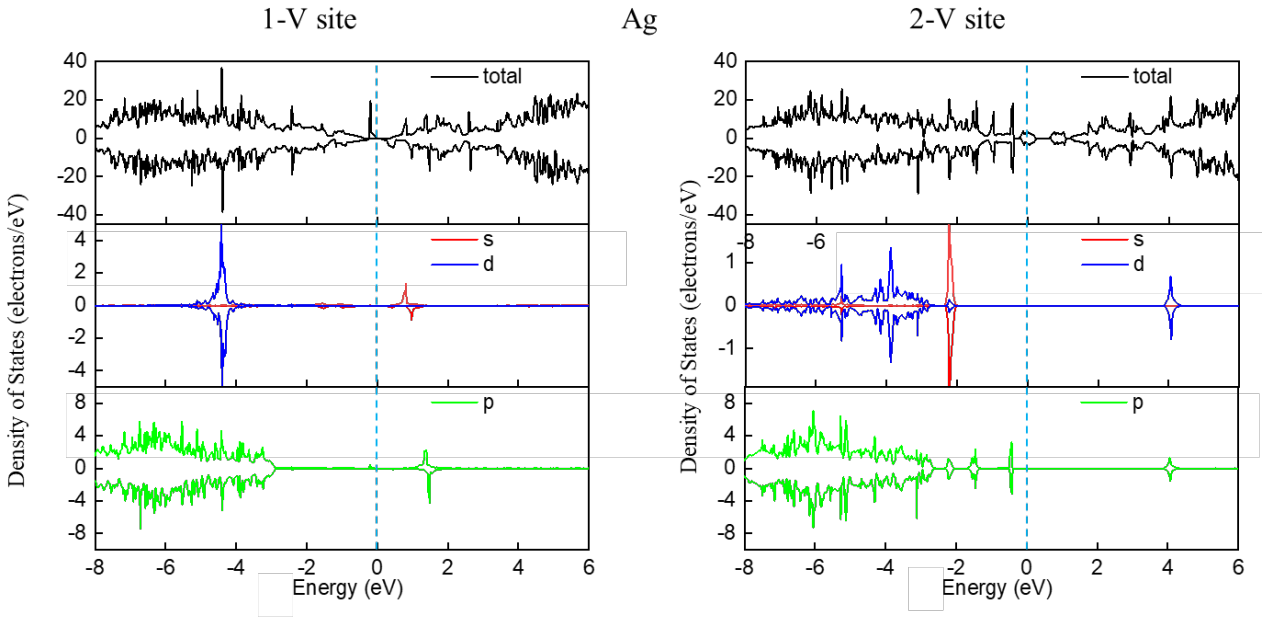


Fig. S2 The DOS and PDOS for Ag on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

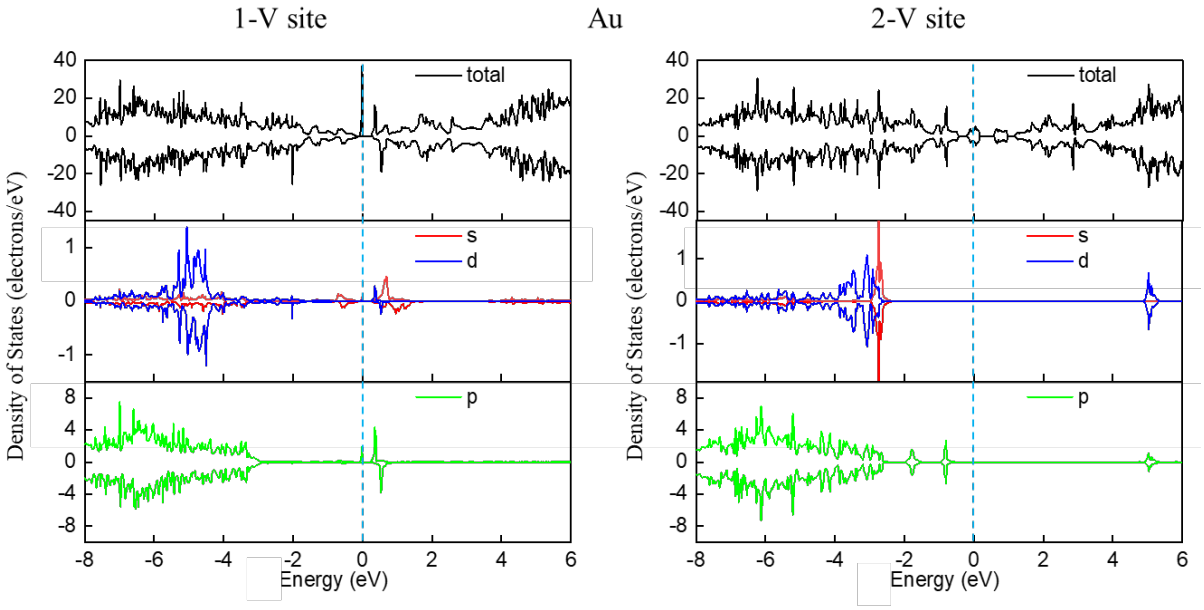


Fig. S3 The DOS and PDOS for Au on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

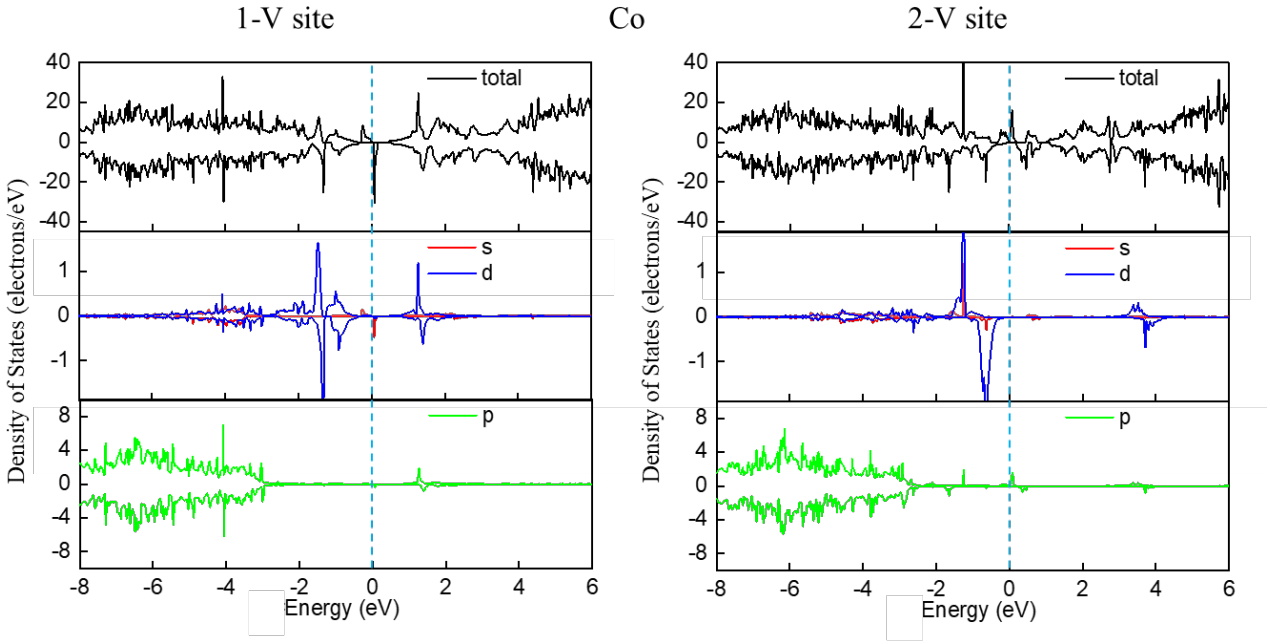


Fig. S4 The DOS and PDOS for Co on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

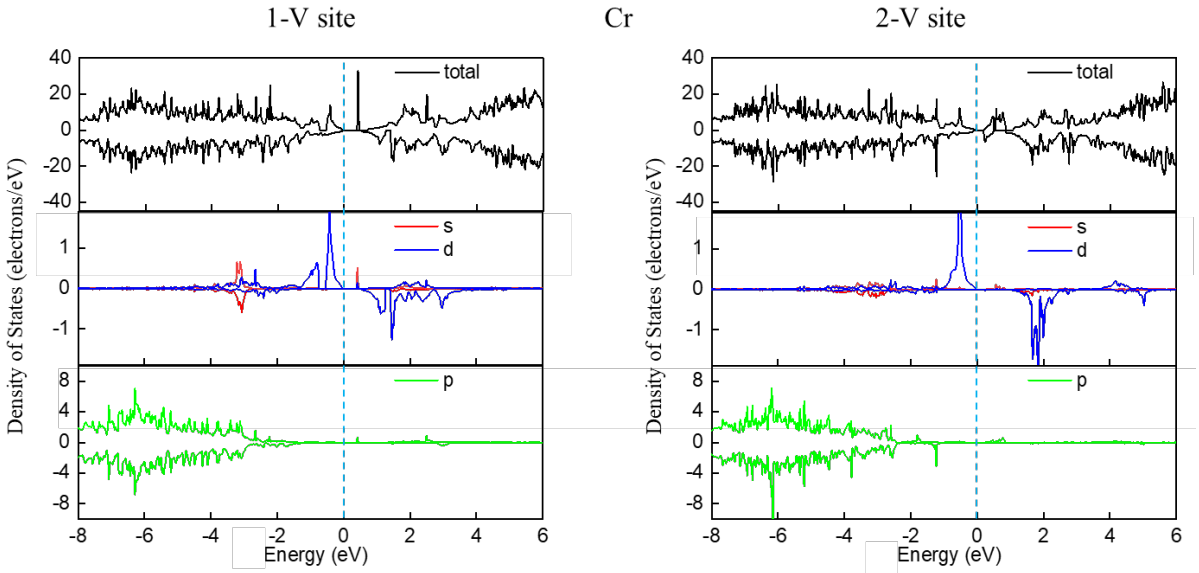


Fig. S5 The DOS and PDOS for Cr on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

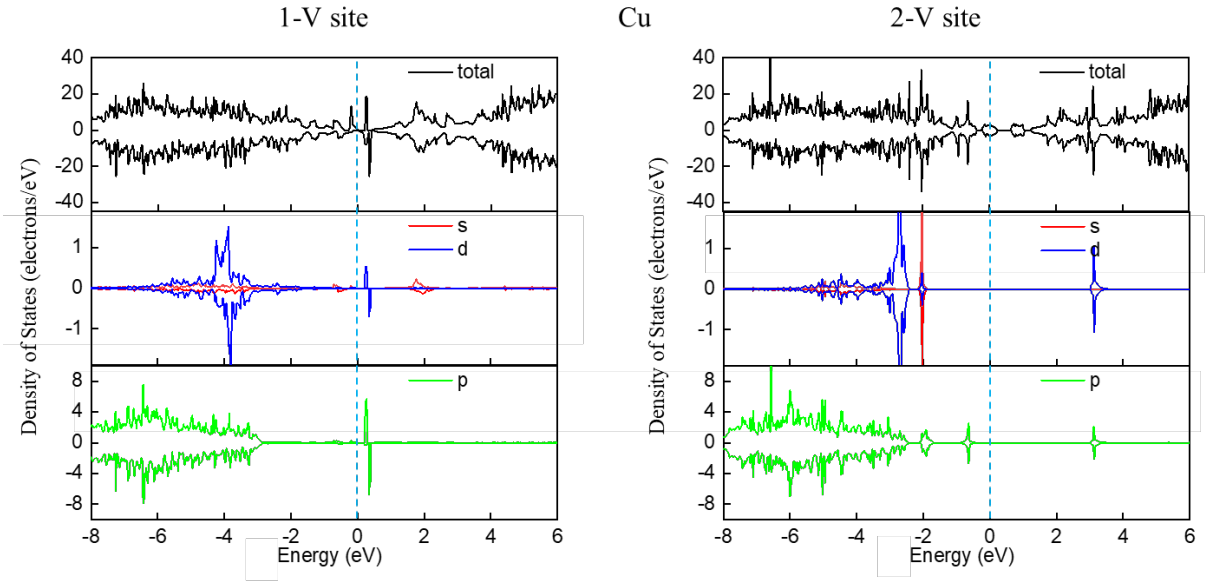


Fig. S6 The DOS and PDOS for Cu on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

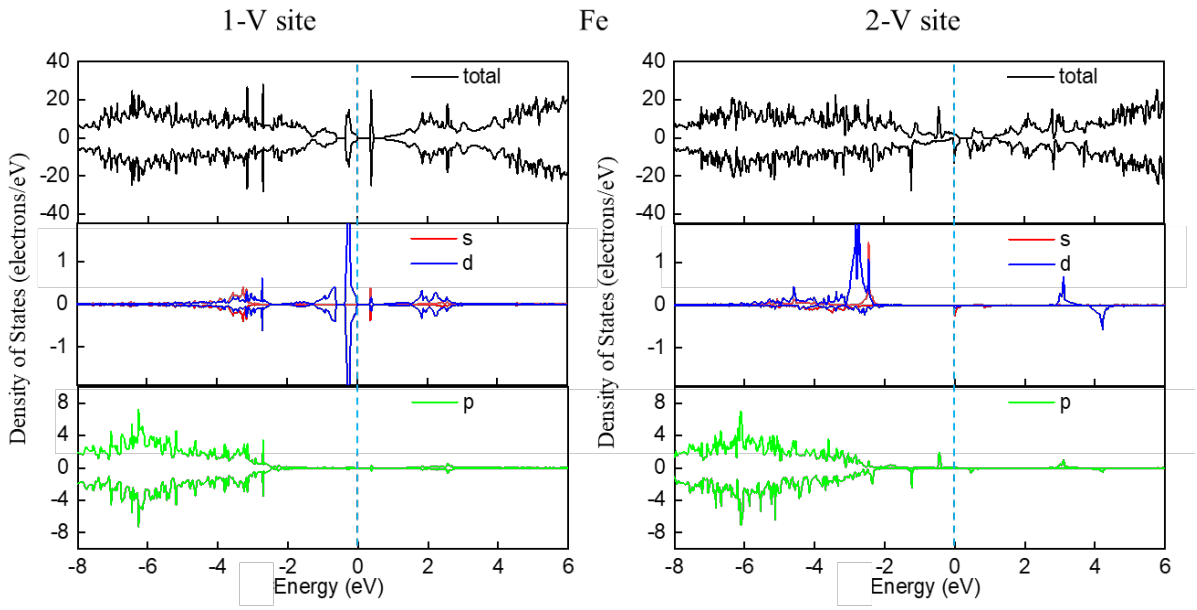


Fig. S7 The DOS and PDOS for Fe on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

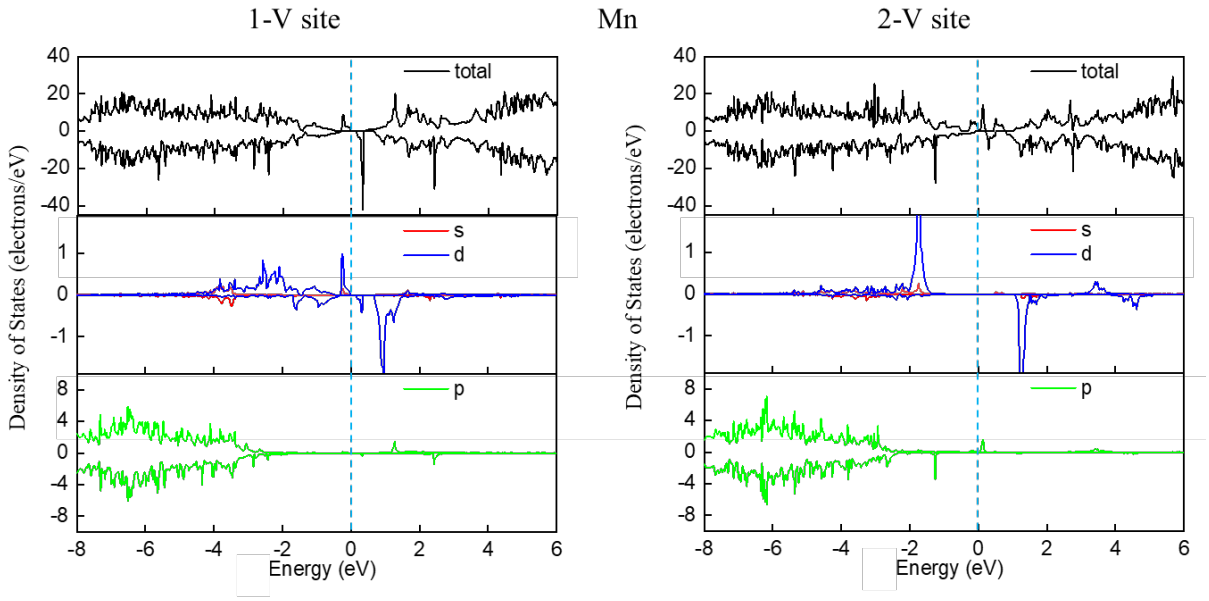


Fig. S8 The DOS and PDOS for Mn on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

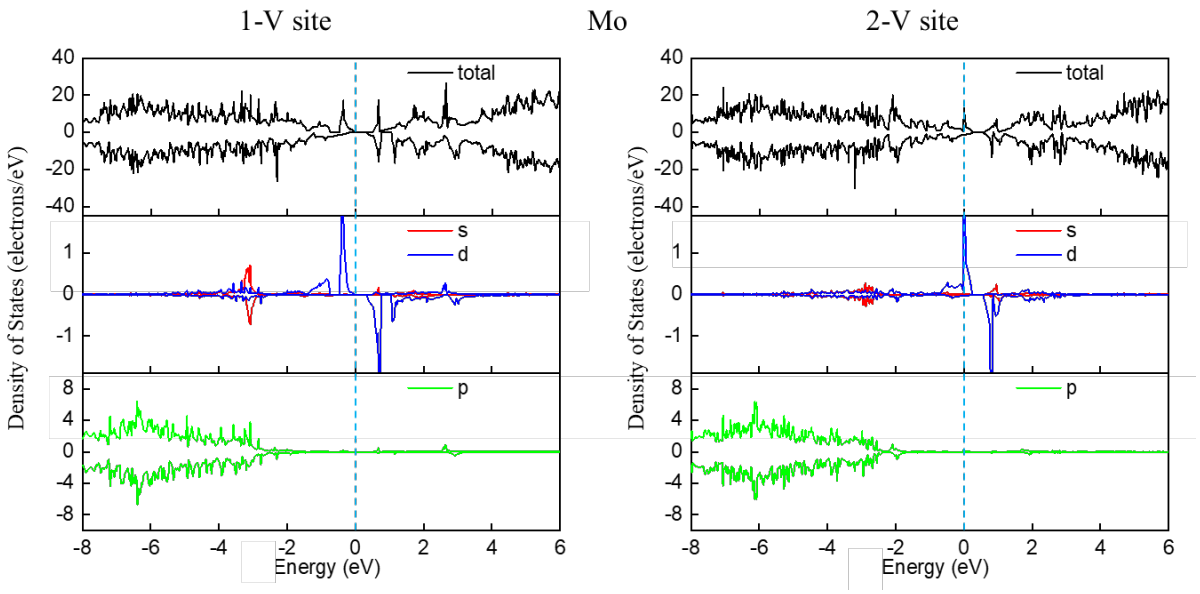


Fig. S9 The DOS and PDOS for Mo on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

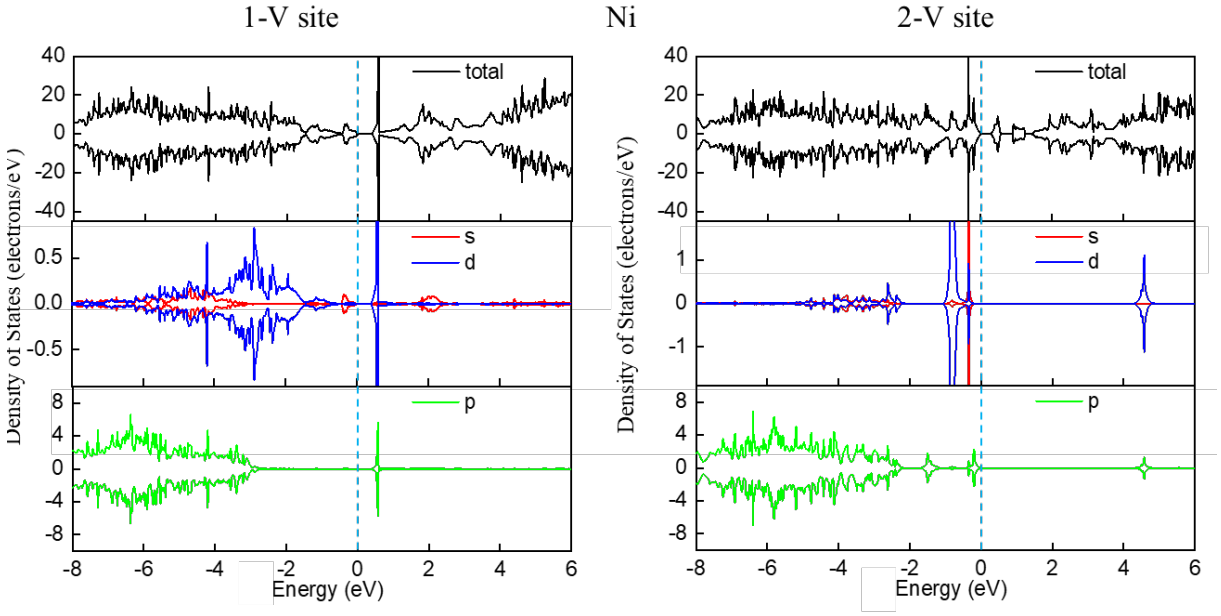


Fig. S10 The DOS and PDOS for Ni on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

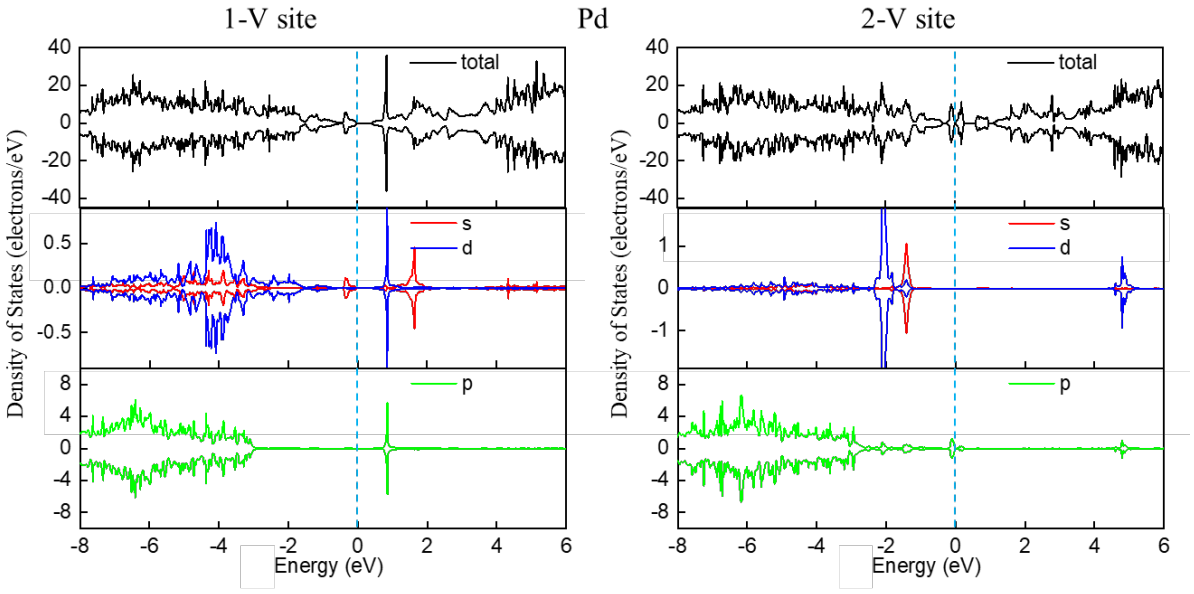


Fig. S11 The DOS and PDOS for Pd on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

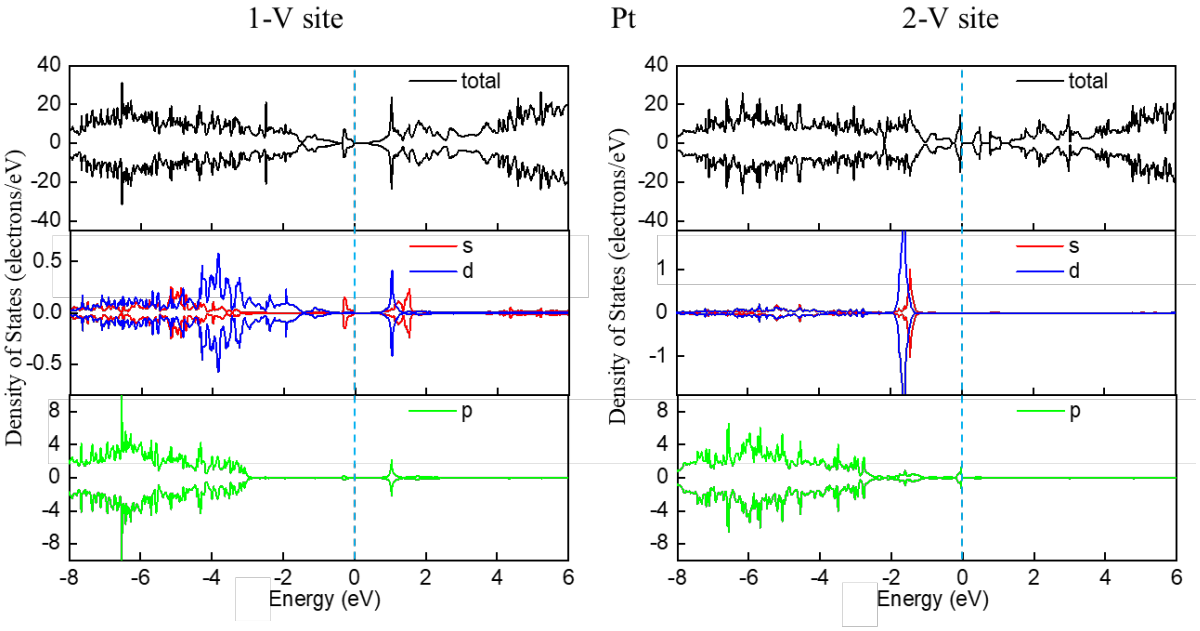


Fig. S12 The DOS and PDOS for Pt on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

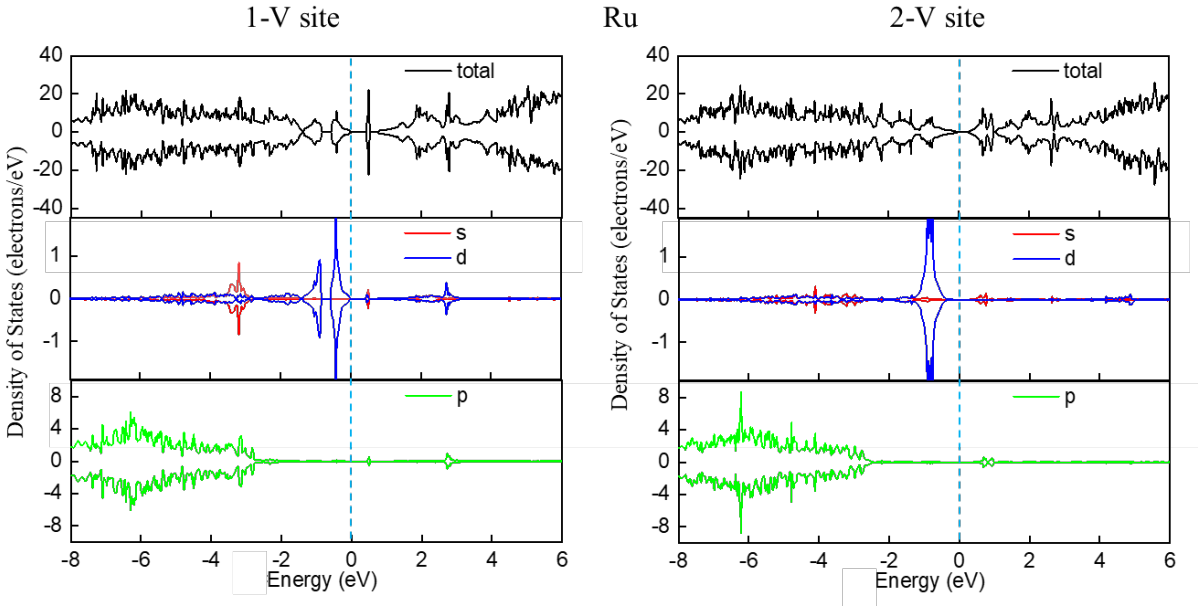


Fig. S13 The DOS and PDOS for Ru on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

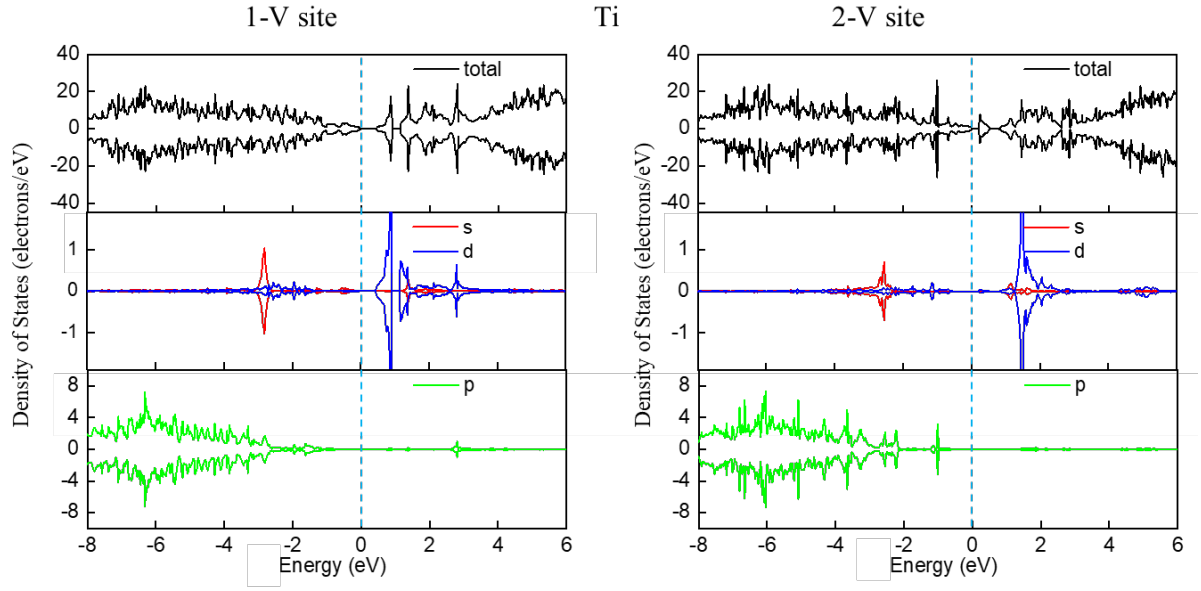


Fig. S14 The DOS and PDOS for Ti on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

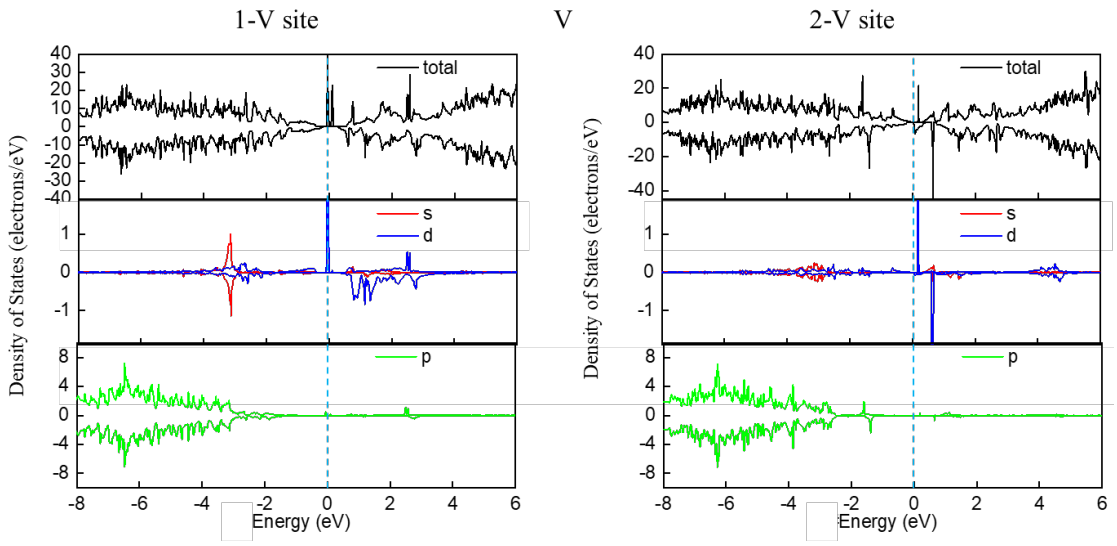


Fig. S15 The DOS and PDOS for V on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1.

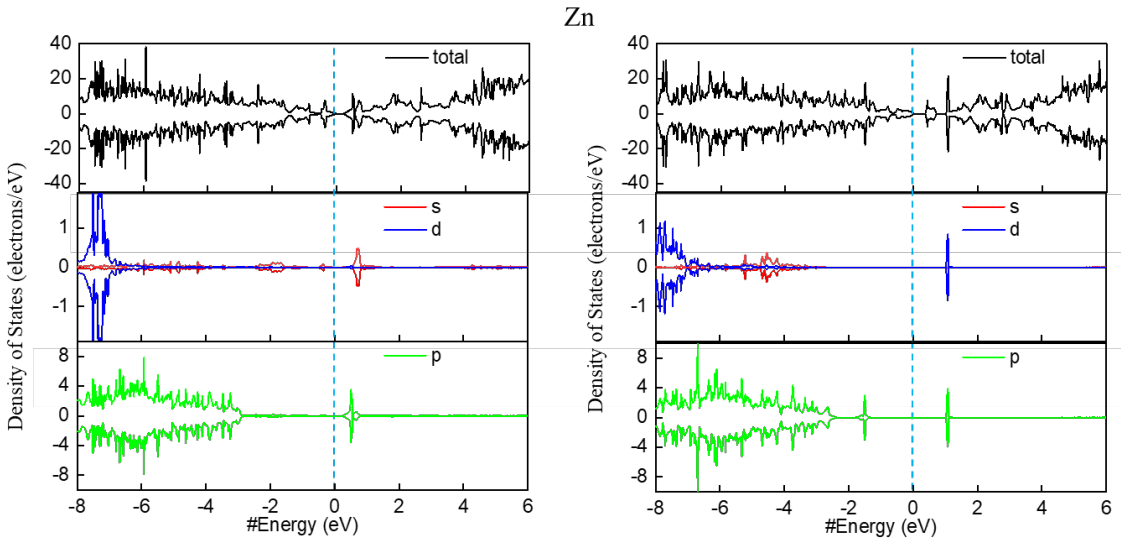


Fig. S16 The DOS and PDOS for Zn on 1-V site and 2-V site. Three plots from top to bottom stand for the total DOS of the system, the PDOS of the metal and the PDOS of the carbon atom, respectively. Related to Figure 1. The polarized charge mainly results from *s*-electron and *d*-electron of the metal atom, mostly concentrated around the metal atom rather than the carbon atoms near the metal atom. An important common character in the DOS for these single atoms on carbon surface is the strong hybridization between the d states of the adatoms and the p state of graphene. The *d* valence states of the metal atom are well below the Fermi level and interact strongly with graphene. The occupation of the spin-up and spin-down states is uneven for Ag, Au, Co, Cr, Cu, Mn, Mo, V on SV-site and Ag, Co, Cr, Cu, Fe, Mn, Mo, V on DV-site of carbon surface, suggesting the exist of magnetic moments. For the Fe, Ni, Pd, Pt, Ru, Ti on SV-site and Au, Ni, Pd, Pt, Ru, Ti, Zn on DV-site of carbon surface, the occupation of the spin-polarized d states is symmetric, resulting in a zero magnetic moment.

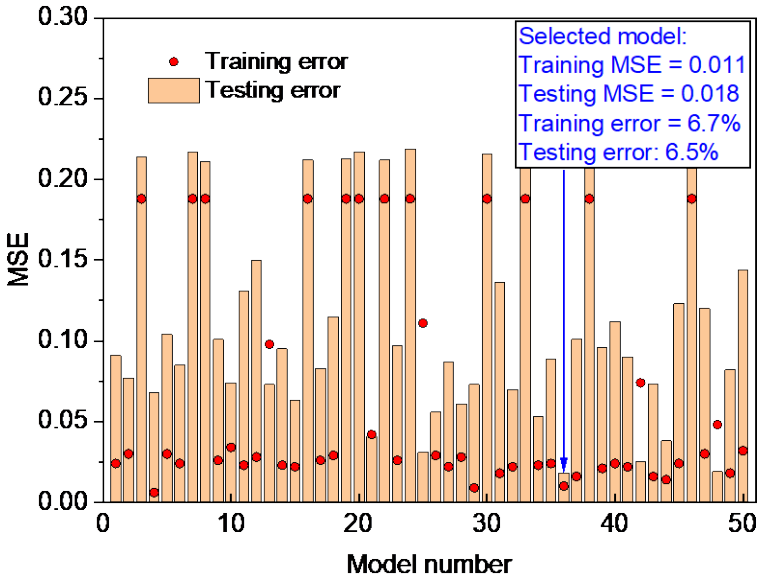


Figure S17 MSE values of the 50 models for the training/testing dataset. Related to Figure 3.

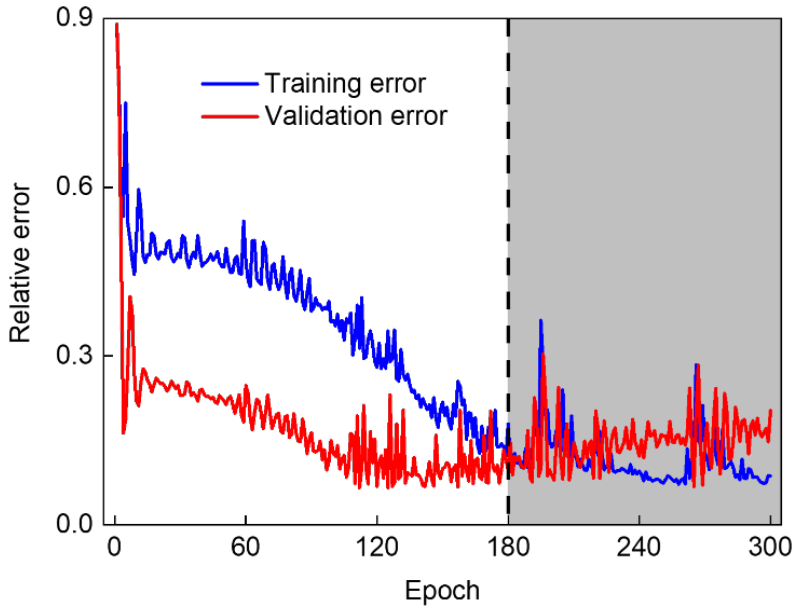


Figure S18 Relative error in training/testing dataset during the training progress. Related to Figure 3.

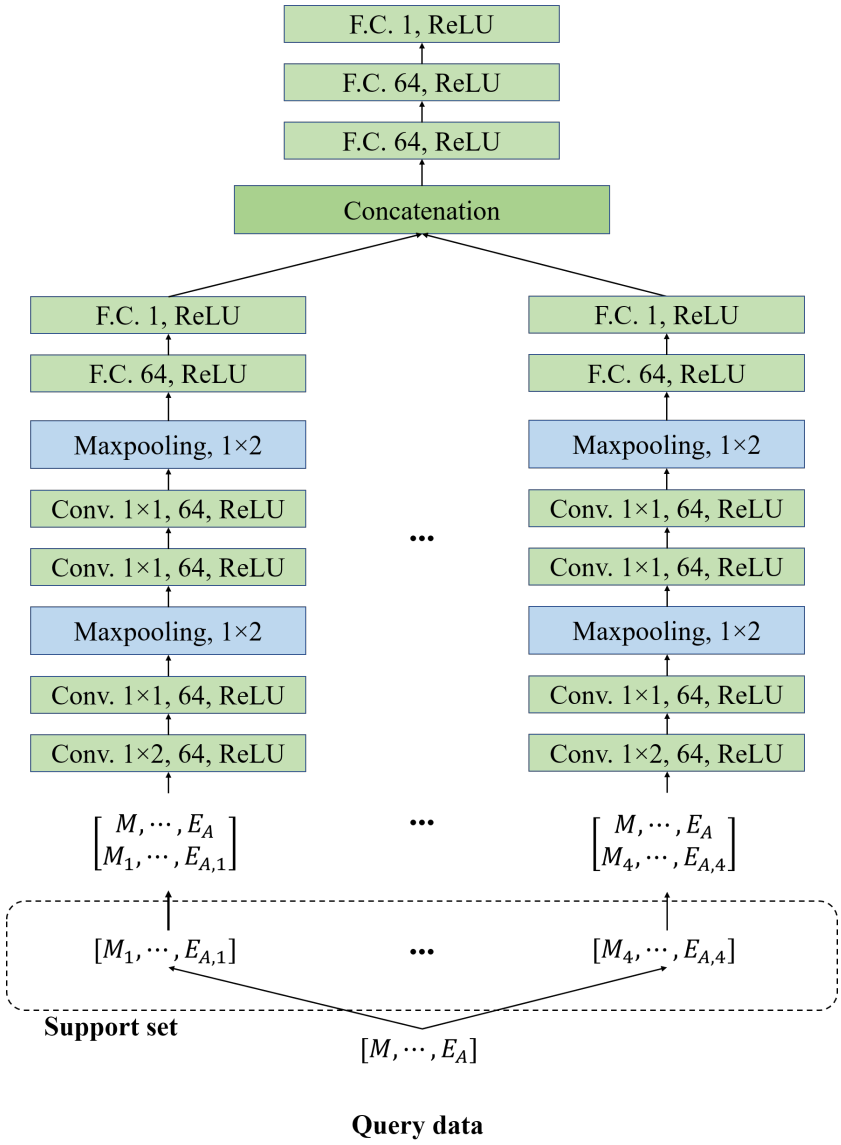


Figure S19 Detailed description of the architecture of the ML model. Related to Figure 2.

Supplemental References

Blöchl, P.E., Jepsen, O., and Andersen, O.K. (1994). Improved tetrahedron method for Brillouin-zone integrations. Physical Review B *49*, 16223-16233.

Kresse, G., and Hafner, J. (1993). Ab initio molecular dynamics for liquid metals. Physical Review B *47*, 558-561.

Kresse, G., and Hafner, J. (1994). Ab initio molecular-dynamics simulation of the liquid-metal--amorphous-semiconductor transition in germanium. Physical Review B *49*, 14251-14269.

Kresse, G., and Joubert, D. (1999). From ultrasoft pseudopotentials to the projector augmented-wave method. Phys Rev B *59*, 1758-1775.

Perdew, J.P., Burke, K., and Ernzerhof, M. (1996). Generalized Gradient Approximation Made Simple. Physical Review Letters *77*, 3865-3868.

Zuluaga, S., and Stolbov, S. (2011). Factors controlling the energetics of the oxygen reduction reaction on the Pd-Co electro-catalysts: Insight from first principles. The Journal of Chemical Physics *135*, 134702.