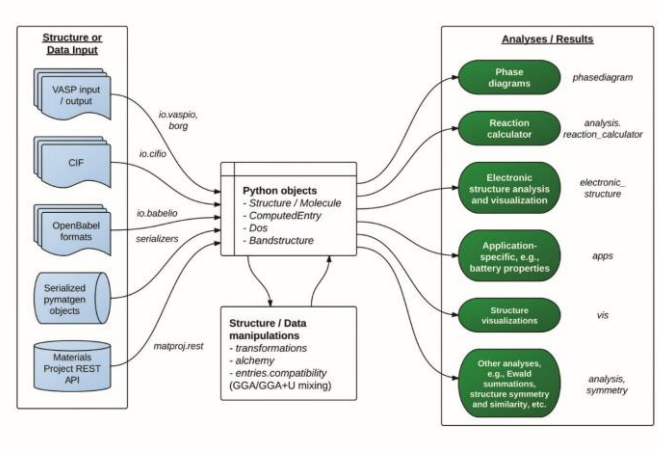
Screening design of HER catalyst based on pre-train model

Chengrui Xu，Haibo Wang，Jingxuan Wang，Tianyi Fang，Huizhen Ni

（Tanwei college，Department of Chemical Engineering, Tsinghua University，Beijing100084）

Abstract



As a kind of good green energy, hydrogen has the advantages of high calorific value and low emission. In order to produce hydrogen efficiently and safely, catalytic electrolysis hydrogen (HER) is currently the most ideal method. On the basis of previous studies, Pt-Co alloy and Ni-P alloy with good catalytic performance as the screening raw materials were selected. Pymatgen was applied to cut the crystals and form the slabs, and the pre-trained large model Equiformer V2 was used to calculate the free energy at the crystal surface adsorption site. Finally, after data processing and analysis, the most ideal crystal and its cutting mode are obtained, namely: CoPt3_mp-1226089_111_1 ,NiP4_mp-769108_110_0 and Ni2P_mp-21167_001_0_5_opt.cif.

Keywords: Python; HER; Deep Learning; catalyst; DFT calculation

1 Introduction

Hydrogen energy is a kind of secondary energy that is rich in sources, green and low carbon and widely used, which could alleviate the climate change issues and energy crisis arising out of the use of non-renewable resources (Holladay et al., 2009). Nowadays, electrochemical water splitting for hydrogen evolution reaction (HER) plays an important role in high-purity hydrogen generation, and its industrial application highly relies on HER electrocatalysts' performance(Wang et al., 2019). In previous studies, Nickel-phosphorus alloy(Moon et al., 2015) and platinum-cobalt alloy(Yang et al., 2022) are verified to have good catalysis activity in HER. According to the data

obtained from the Material Project, there are 5341 kinds of nickel-phosphorus alloy and 4233 kinds of platinum-cobalt alloy adsorption structures worth exploring. However, Due to the complex surface chemistry, screening them either via first principles calculations or experiments are limited by their intense resource requirements. In such a scenario, using machine learning models to assist in exploring the performance of several promising HER catalysts will be an advisable choice, which is expected to accelerate potential electrocatalysts discovery by a factor of at least 10(Toyao et al., 2019).

Here, we have used GEMNET-DT, a pretrained graph neural network model to compute the activity of Ni-P alloy and Pt-Co alloy catalysts. GEMNET-DT is a variant of geometric message passing neural networks, which are GNNs with spherical representations(Gasteiger et al., 2021). It is a universal approximator for predictions that are invariant to translation, and equivariant to permutation and rotation. Figure 1 shows its structure. It can predict the trajectories, energies, and forces with initial adsorbate surface geometries. Thus, adsorption energies of H* for different catalysts can be obtained from GEMNET-DT. These values can be used to compute the activity of catalysts in a quasi-equilibrium microkinetic model(D. Wang et al., 2023).

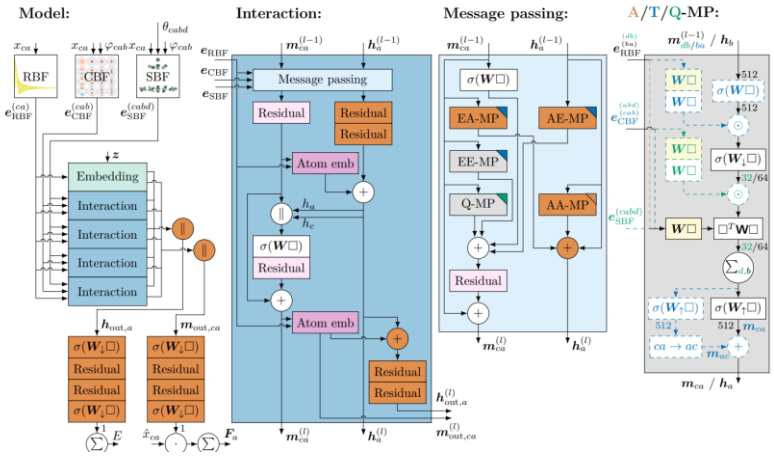


Figure 1 GEMNET-DT's structure

2 Methods

A. Data Preparation and Representation

14 thermodynamically stable (energy above hull < 0.1) Ni_xP_y and 6 thermodynamically stable Co_xPt_y bulks were obtained from database Material Project, and the slab models were cleaved

from each structure by applying the pymatgen package from Python. Figure 2 illustrates the process of how a particular adsorption geometry was obtained through cleaving, adsorption points forming and hydrogen atom attachment.

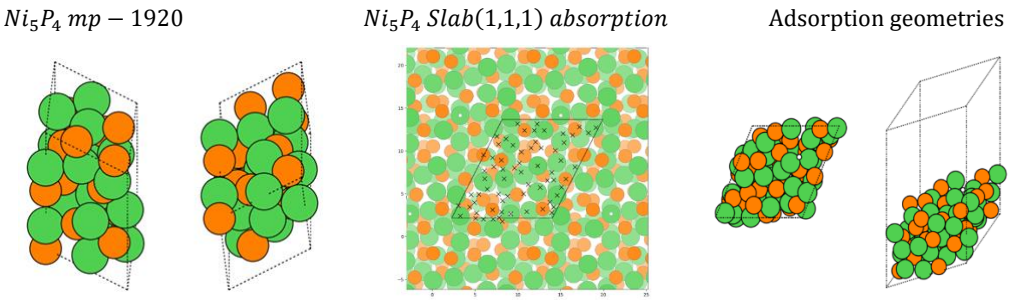


Figure 2 Illustration of the creation of slab-adsorbate structures for various intermediates.

B. Module Selection

GEMNET-DT model, pretrained on the large OC20 dataset (Chanussot et al., 2021), was used to compute adsorbate binding energies. The predecessor in this field has well proved the reliability of this module. In Gasteiger and his coworkers' work, GEMNET-OC module was proved to be much more powerful and accurate in stimulating DFT calculation compared to GEMNET-DT module (Gasteiger et al., 2022).

C. Computation of Adsorption Free Energies

After successfully creating a slab module for the surface and placing an adsorbate on it as an initial guess, the relaxation of the lowest energy geometry can be performed. Adsorption free energies were then obtained as a sum of the GEMNET-DT adsorption energy and the vibrational contributions. These computations were driven using the thermochemistry module of ASE (https://wiki.fysik.dtu.dk/ase/ase/thermochemistry/thermochemistry.html).

D. Module Tuning and DFT Calculations

To further prove that the module applied is reliable, DFT calculation is a crucial step. In this step, the number of layers, effects of relaxation and unit cell size etc. shall be in consideration and finally formula1 is applied to generate the accurate value (specifically speaking, VASP was the

main tool applied). The given results shall be fed back to the original calculation module as a test group thus further tuning it theoretically.

$$\hat{H}\Psi = \left[\hat{T}+\hat{V}+\hat{U}\right]\Psi = \left[\sum_i^N\left(-\frac{\hbar^2}{2m_i}\nabla_i^2\right)+\sum_i^N V(r_i)+\sum_{i<j}^N U(r_i,r_j)\right]\Psi = E\Psi$$

Formula 1 Schrodinger's equation

3 Results

GEMNET-DT is used to predict the adsorption energy of hydrogen. In all possible alloy structures of nickel-phosphorus alloy and cobalt-platinum alloy, the section mode with low adsorption energy and the adsorption site of hydrogen atom were selected and visualized. Part of the results were shown in Figure 3.

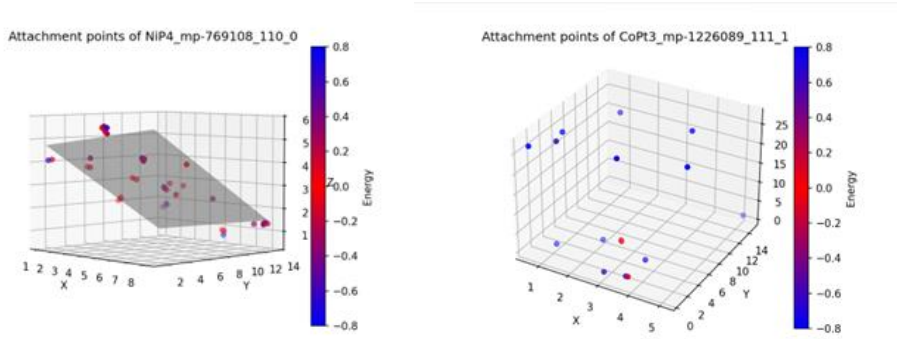


Figure 3 Attachment points

Then, we analyzed the catalyst properties of various alloy sections by comparing the energy changes before and after hydrogen adsorption. Among all possible alloy structures, the hydrogen atom adsorption energy of Ni2P and CoPt3 alloy sections is -0.01eV and -0.03eV respectively, indicating relatively good catalyst performance (see Figure 4).

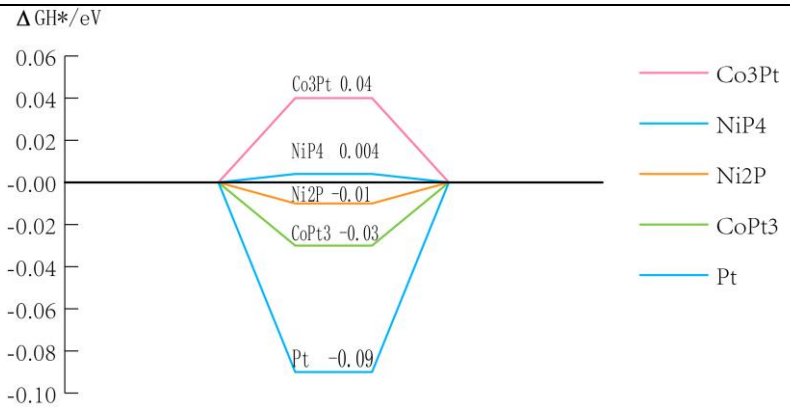


Figure 4 hydrogen atom adsorption energy

However, when we used the crystal data for DFT calculations, the data results of some crystal sections were found to be divergent, that is, these crystal interfaces are not actually available. Therefore, we conducted the secondary screening of the obtained crystal structure, removing some of the inaccessible crystal sections and some of the adsorption sites whose adsorption energy was too small and thus the calculation error was too great. Figure 5 shows the DFT calculation of one slab example.

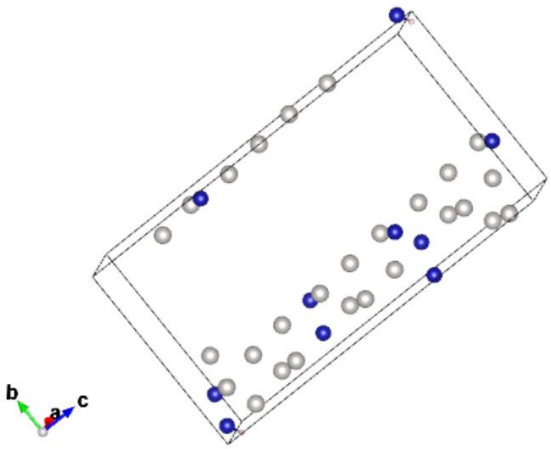


Figure 5 DFT calculation

To further adjust the model parameters, we keep the number of layers at 6, fix the bottom 5 layers for optimization, and apply wide wrapping to the crystals. Considering the poor adaptability of the GEMNET model in hydrogen atom adsorption, the model was adjusted to EquiformerV2-83M-S2EF-OC20-2M. At the same time, we plotted the plot of the unit cell slice number and the

adsorption energy of each hydrogen atom, and used the clustering algorithm to find out the relationship between the Miller index and the adsorption energy of hydrogen atoms.

Due to the lack of computing power of the personal computer, we were unable to perform further training on all the screened crystals. Figure 6 shows the box plots of energy values for different Miller indices of Ni_2P.

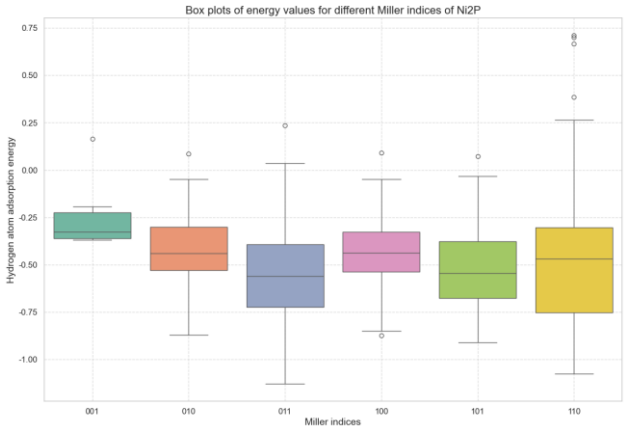


Figure 6 the box plots of energy values for different Miller indices of Ni_2P.

From the box plot, it can be seen that Ni_2P crystal has a small adsorption energy and a more stable adsorption structure in the (0,0,1) section, and the energy variance of each adsorption site is small, which is an ideal catalyst for HER reaction. Finally, we calculated the Ni2P and obtained its most promising sections, which were stored as follows:

The Ni2P file with the smallest hydrogen atom adsorption energy value is:

Ni2P_mp-21167_001_0_5_opt.cif, and the energy value is: -0.30458984375

In the screened crystal section, the Ni2P, especially for (0 0 1) surface, exhibited high electron density in d-band extending to near Fermi level, resulting in the strong interaction with hydrogen. As for Pt-Co alloys, the hydrogen adsorption energy of CoPt3 is -0.12eV, which is the smallest among all Pt-Co alloy crystal structures (see Figure 5).

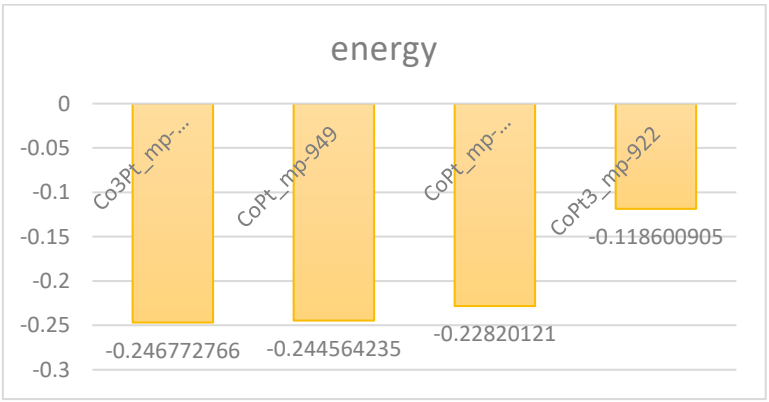


Figure 5 Some Co-Pt alloy's hydrogen adsorption energy

The results of both crystals are consistent with the previous literature data, which proves the reliability of this prediction method(Moon et al., 2015)(Y. Wang et al., 2022).

4 Conclusion and Discussion

In the present study, we have screened Nickel-phosphorus and platinum-cobalt alloys to identify efficient catalysts for HER. Using the GEMNET-DT model, we find Ni_2P surfaces (0 0 1) and Pt_3Co surfaces (1 1 0) have good catalyst activity, whose hydrogen adsorption energies are respectively -0.01 and -0.24. For most of our catalysts, especially the Pt_3Co, its surfaces (1 1 0)'s hydrogen adsorption energy is very close to the data in the preview literature(Y. Wang et al., 2022). It proves that the GEMNET-DT model has a reference value in predicting the performance of catalysts.

However, because the accuracy of the GEMNET-DT model and the prediction field involved is still limited, the use of the model still needs to be combined with experimental and theoretical calculations. To refine our research, the following work will be carried out.

1. Further discuss the stability of the obtained structure. Using TEM, XAS, XRD, ICP-MS, Raman, and other characterizations, to probe the actual active center and monitor the structural evolution of the catalysts during operation. This index is particularly important when measuring the actual use value of the catalyst.

2. Calculated turnover frequency (TOF). It represents the number of catalytic reactions or the number of target products generated or the number of reactants consumed in unit time and at the unit active site under a certain temperature, pressure, reactant proportion and a certain reaction degree. It can more directly reflect the performance of the catalyst.

3. Model optimization and improvement. Some of the predicted hydrogen adsorption energies are far less than the error range, causing them to fail to reflect the true hydrogen adsorption energies. Some crystal sections were found to be divergent for DFT. To solve this problem, it is expected that the input crystal parameters can be adjusted to optimize the model operations.

Acknowledgements

The authors are grateful to Prof. Wang, Asst. Honghao Chen, Siwei Fu, Kuan liu and Friend Jinyu Yang for their help.

Group Credits Statement:

选题与文献阅读：在文献搜索与阅读阶段，我们小组每位同学分工查找了不同阶段，不同方向的文献，组长制作了共享文档，小组按照文献名称——搜集人——内容梗概——目的——方法——结论——链接进行了汇总，现将共享文档链接附录于下：

【金山文档 | WPS 云文档】
https://kdocs.cn/l/cseU1yF4PiQp

接着我们进行了环境的配置和代码框架的搭建，该部分由许呈睿和王海博完成，我们搭建了完整的高通量催化剂筛选与预训练模型运用与调试框架，并能够实现其与各类晶体文件格式的转换与读取。

在汇报展示阶段，小组成员在组长的带领下分工完成了 PPT 的制作，最终由许呈睿，倪汇振，方天翼同学汇报展示。特别感谢杨谨毓同学提供的理论指导与计算帮助！
详细分工如下：
许呈睿：数据收集和预处理，机器学习模型开发，数据筛选，PPT，论文修改与整合
方天翼：DFT 计算，汇报，理论指导
王海博：数据收集和预处理，机器学习模型开发，论文 results 部分
王景煊：数据可视化，PPT，论文 methods 部分，DFT 计算
倪汇振：数据可视化，PPT 论文 Introduction 和 conclusion&discussion 部分

References:

[1] Chausson, L., Das, A., Goyal, S., Lavril, T., Shuaibi, M., Riviere, M., Tran, K., Heras-Domingo, J., Ho, C., Hu, W., Palizhati, A., Sriram, A., Wood, B., Yoon, J., Parikh, D., Zitnick, C. L., & Ulissi, Z. (2021). Open Catalyst 2020 (OC20) Dataset and community challenges. *ACS Catalysis*, *11*(10), 6059–6072. https://doi.org/10.1021/acscatal.0c04525

[2] Gasteiger, J., Becker, F., & Günnemann, S. (2021, June 2). *GEMNET: Universal Directional Graph Neural Networks for Molecules*. arXiv.org. https://arxiv.org/abs/2106.08903

[3] Gasteiger, J., Shuaibi, M., Sriram, A., Günnemann, S., Ulissi, Z., Zitnick, C. L., & Das, A. (2022). GEMNET-OC: Developing graph neural networks for large and diverse molecular simulation datasets. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2204.02782

[4] Holladay, J., Hu, J., King, D., & Wang, Y. (2009). An overview of hydrogen production technologies. *Catalysis Today*, *139*(4), 244–260. https://doi.org/10.1016/j.cattod.2008.08.039

[5] Moon, J., Jang, J., Kim, E., Chung, Y., Yoo, S. J., & Lee, Y. (2015). The nature of active sites of Ni2P electrocatalyst for hydrogen evolution reaction. *Journal of Catalysis*, *326*, 92–99. https://doi.org/10.1016/j.jcat.2015.03.012

[6] Toyao, T., Maeno, Z., Takakusagi, S., Kamachi, T., Takigawa, I., & Shimizu, K. (2019). Machine Learning for Catalysis Informatics: Recent applications and Prospects. *ACS Catalysis*, *10*(3), 2260–2297. https://doi.org/10.1021/acscatal.9b04186

[7] Wander, B., Broderick, K., & Ulissi, Z. W. (2022, August 26). *Catlas: an automated framework for catalyst discovery demonstrated for direct syngas conversion*. arXiv.org. https://arxiv.org/abs/2208.12717

[8] Wang, D., Jin, L., Liu, M., Lee, T. G., Peera, S. G., & Liu, C. (2023). The graphene-supported Lanthanum oxide cluster as efficient bifunctional electrocatalyst for oxygen reaction. *Molecular Catalysis*, *535*, 112879. https://doi.org/10.1016/j.mcat.2022.112879

[9] Wang, M., Zhang, L., Huang, M., Zhang, Q., Zhao, X., He, Y., Lin, S., Pan, J., & Zhu, H. (2019). One-step synthesis of a hierarchical self-supported WS2 film for efficient electrocatalytic hydrogen evolution. *Journal of Materials Chemistry A*, *7*(39), 22405–22411. https://doi.org/10.1039/c9ta07868a

[10] Wang, Y., Wu, W., Chen, R., Lin, C., Mu, S., & Cheng, N. (2022). Reduced water dissociation barrier on constructing Pt-Co/CoOx interface for alkaline hydrogen evolution. *Nano Research*, *15*(6), 4958–4964. https://doi.org/10.1007/s12274-022-4128-6

[11] Yang, W., Cheng, P., Li, Z., Lin, Y., Li, M., Zi, J., Shi, H., Li, G., Lian, Z., & Li, H. (2022). Tuning the Cobalt–Platinum alloy regulating Single-Atom platinum for highly efficient hydrogen evolution reaction. *Advanced Functional Materials*, *32*(39). https://doi.org/10.1002/adfm.202205920