

DOI: <https://doi.org/10.52714/dthu.14.04S.2025.1544>

ỨNG DỤNG MÔ HÌNH MÁY HỌC TRONG MÔ PHỎNG VẬT LIỆU BẰNG PYTHON TRÊN GOOGLE COLAB

Nguyễn Hải Đăng^{1,2*}, Hồ Nhã Nghi³, Đặng Minh Triết⁴, và Nguyễn Thành Tiên²

¹Khoa Cơ bản, Trường Đại học Nam Cần Thơ, Việt Nam

²Khoa Khoa học Tự nhiên, Trường Đại học Cần Thơ, Việt Nam

³Trường THPT Chuyên Huỳnh Mãn Đạt, Kiên Giang, Việt Nam

⁴Trường Sư phạm, Trường Đại học Cần Thơ, Việt Nam

*Tác giả liên hệ, Email: nhdang@nctu.edu.vn

Lịch sử bài báo

Ngày nhận: 03/4/2025; Ngày nhận chỉnh sửa: 15/5/2025; Ngày duyệt đăng: 19/5/2025

Tóm tắt

Nghiên cứu này xây dựng mô hình máy học sử dụng mạng neural nhân tạo để huấn luyện thể tương tác nguyên tử, dựa trên nền tảng máy học từ dữ liệu bởi mô phỏng động lực học phân tử. Mục tiêu là dự đoán năng lượng hình thành của graphene nanoribbon cạnh armchair khuyết một nguyên tử carbon. Thuật toán học sâu được triển khai trên nền tảng Google Colab bằng ngôn ngữ Python.

Từ khóa: Google Colab, máy học, học sâu, mạng neural, Python.

Trích dẫn: Nguyễn, H. Đ., Hồ, N. N., Đặng, M. T., & Nguyễn, T. T. (2025). Ứng dụng mô hình máy học trong mô phỏng vật liệu bằng Python trên Google Colab. *Tạp chí Khoa học Đại học Đồng Tháp*, 14(04S), 27-35. <https://doi.org/10.52714/dthu.14.04S.2025.1544>

Copyright © 2025 The author(s). This work is licensed under a CC BY-NC 4.0 License.

**APPLICATION OF MACHINE LEARNING MODEL IN MATERIAL
SIMULATION USING PYTHON ON GOOGLE COLAB**

Nguyen Hai Dang^{1,2*}, Ho Nha Nghi³, Dang Minh Triet⁴, and Nguyen Thanh Tien²

¹*Faculty of Basic Sciences, Nam Can Tho University, Vietnam*

²*College of Natural Sciences, Can Tho University, Vietnam*

³*Huynh Man Dat Gifted High School, Kien Giang, Vietnam*

⁴*School of Education, Can Tho University, Vietnam*

** Corresponding author, Email: nhdang@nctu.edu.vn*

Article history

Received: 03/4/2025; Received in revised form: 15/5/2025; Accepted: 19/5/2025

Abstract

This study builds a machine learning model using artificial neural networks to train machine learning-based atomic interaction potentials from data by molecular dynamics simulations. The goal is to predict the formation energy of armchair-edge graphene nanoribbon with a missing carbon atom. The deep learning algorithm is implemented on the Google Colab platform using Python language.

Từ khóa: *Deep learning, Google Colab, machine learning, neural networks, Python.*

1. Giới thiệu

Trong bối cảnh khoa học vật liệu hiện đại, việc kết hợp máy học (Machine Learning – ML) với các phương pháp mô phỏng truyền thống đang mở ra nhiều hướng tiếp cận mới trong thiết kế và khám phá vật liệu. Máy học không chỉ giúp giảm chi phí tính toán mà còn nâng cao hiệu quả dự đoán các tính chất vật liệu một cách chính xác, đặc biệt khi được huấn luyện với dữ liệu từ các mô phỏng lý thuyết đáng tin cậy như động lực học phân tử (Molecule Dynamic - MD) và lý thuyết phiếm hàm mật độ (Density Functional Theory - DFT) (Butler, 2018; Schmidt, 2019). Một trong những ứng dụng nổi bật của máy học trong vật lý vật liệu là xây dựng các thế tương tác nguyên tử máy học – Machine Learning Interatomic Potentials (MLIPs). Các mô hình MLIP, điển hình như mạng neural nhân tạo (Artificial Neural Networks – ANN) và Moment Tensor Potentials (MTP), có khả năng mô tả chính xác thế năng lượng của các hệ vật liệu phức tạp, với độ chính xác gần tiệm cận DFT nhưng tốc độ tính toán nhanh hơn nhiều lần (Behler, 2016; Shapeev, 2016). Việc huấn luyện các mô hình MLIP thường dựa trên dữ liệu thu được từ mô phỏng DFT hoặc MD, qua đó cho phép mở rộng ứng dụng đến các bài toán động lực học tiêu tốn nhiều tài nguyên tính toán mà các phương pháp ab initio như DFT không thể xử lý trực tiếp.

Với sự phát triển của công nghệ tính toán, việc triển khai các mô hình máy học phục vụ mô phỏng vật liệu đã trở nên dễ tiếp cận hơn nhờ các nền tảng như Google Colab, một môi trường điện toán đám mây hỗ trợ Python, cho phép khai thác sức mạnh tính toán của bộ xử lý đồ họa (GPU) miễn phí và tích hợp nhiều thư viện học máy hiện đại như TensorFlow, PyTorch, và Scikit-learn. Các nghiên cứu gần đây đã chứng minh hiệu quả của việc kết hợp ANN và dữ liệu MD để xây dựng MLIP cho nhiều hệ vật liệu khác nhau, bao gồm silicon, kim loại chuyển tiếp, oxit và vật liệu hai chiều (Nguyen, 2025; Zuo, 2020; Jinnouchi, 2019).

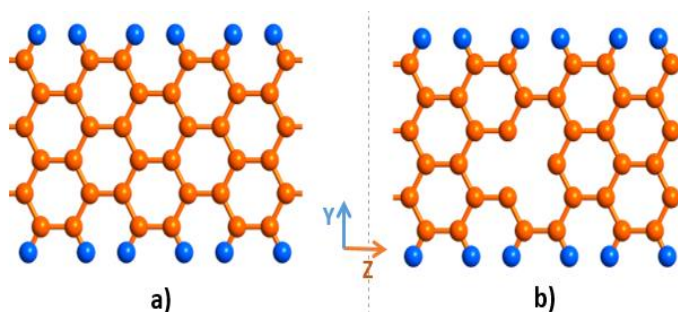
Graphene nanoribbon (GNR) là một vật liệu thu hút sự chú ý lớn nhờ vào các tính chất điện tử, cơ học và hóa học đặc biệt. Tùy vào kiểu cạnh (zigzag hoặc armchair) và sự hiện diện của các khuyết tật, GNR có thể biểu hiện đặc tính kim loại, bán dẫn hoặc thậm chí từ tính, khiến nó trở thành ứng viên tiềm năng cho nhiều ứng dụng trong điện tử nano, cảm biến và vật liệu tiên tiến (Yazyev, 2010; Nguyen, 2018). Các khuyết tật như mất nguyên tử carbon (vacancy defects) trong GNR có thể được tạo ra có chủ đích hoặc phát sinh trong quá trình tổng hợp, đóng vai trò quan trọng trong việc điều chỉnh các tính chất điện tử, cơ học và hóa học của vật liệu. Những khuyết tật này ảnh hưởng trực tiếp đến năng lượng hình thành và độ ổn định cấu trúc của GNR (Zhang, 2022; Zhao, 2022).

Trong nghiên cứu này, chúng tôi xây dựng mô hình máy học sử dụng mạng neural nhân tạo để huấn luyện thế năng lượng MLIP, từ dữ liệu động lực học phân tử cho graphene nanoribbon có cạnh armchair khuyết một nguyên tử carbon. Mô hình được tối ưu nhằm dự đoán năng lượng hình thành của cấu trúc, một đại lượng quan trọng phản ánh độ ổn định nhiệt động của hệ. Các kết quả chủ yếu tập trung vào các chỉ số sai số căn quân phương (RMSE) và hệ số xác định R^2 của các mô hình, qua đó tiến hành so sánh và tìm ra mô hình tối ưu nhất. Toàn bộ quá trình được thực hiện bởi ngôn ngữ Python trên nền tảng Google Colab, với mục tiêu xây dựng một quy trình mô phỏng linh hoạt, có thể mở rộng và tái sử dụng cho các hệ vật liệu khác nhau.

2. Phương pháp nghiên cứu

Ban đầu, mẫu graphene nanoribbon dạng cạnh armchair có chức hóa các nguyên tử hydrogen, được tạo khuyết một nguyên tử carbon (Hình 1), được xây dựng thông qua phần mềm QuantumATK (Smidstrup, 2019). Phần mềm này cho phép tạo ra các mô hình cấu trúc chính xác của graphene nanoribbon, bao gồm việc xác định vị trí và số lượng nguyên tử trong hệ thống. Các thông số cấu trúc của mẫu được sử dụng làm đầu vào cho mô phỏng động lực

học phân tử phiếm hàm mật độ (AIMD), phần mềm VASP được sử dụng để tiến hành tính toán các đặc tính cấu trúc, bao gồm các thông tin về năng lượng, và lực tương tác được lưu trữ trong tệp dữ liệu OUTCAR (Hafner, 2008).



Hình 1. Cấu trúc graphene nanoribbon dạng biên armchair (a) và cấu trúc khuyết một nguyên tử carbon (b). Quả cầu màu cam và màu xanh lần lượt đại diện cho nguyên tử carbon và hydrogen.

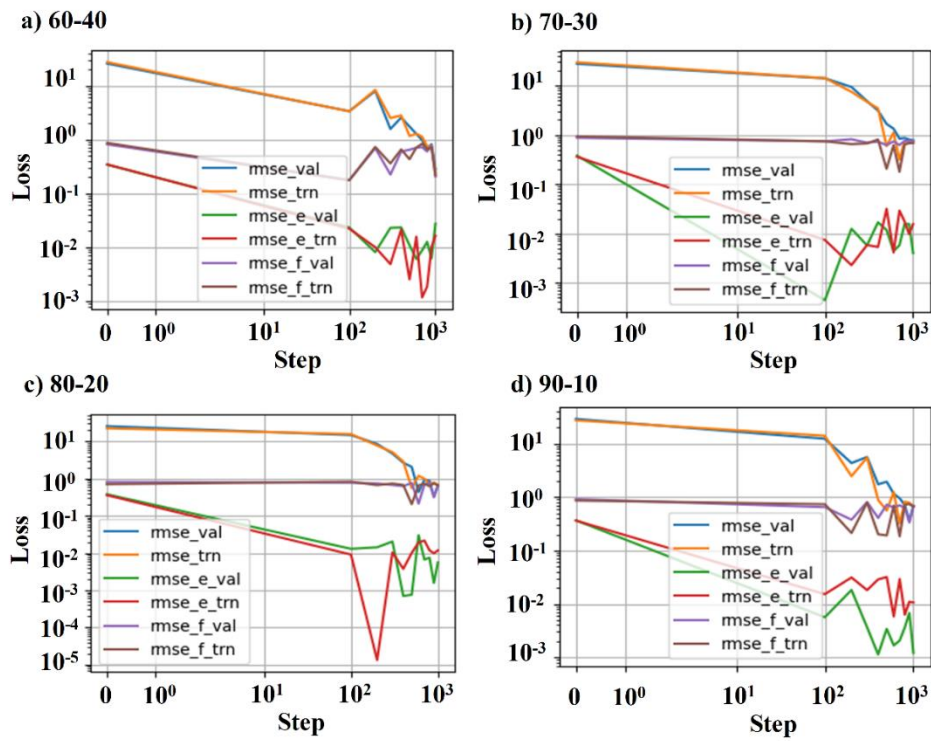
Dựa trên tệp dữ liệu OUTCAR thu được từ các mô phỏng AIMD, bước tiếp theo là huấn luyện thể MLIP bằng DeePMD-kit (Wang, 2018; Zeng, 2025). Quá trình này sử dụng các thông số cấu trúc và năng lượng từ VASP để xây dựng một mô hình tiềm năng giữa các nguyên tử. Việc sử dụng thể MLIP giúp cải thiện hiệu suất tính toán so với phương pháp AIMD truyền thống, đồng thời duy trì độ chính xác trong việc mô phỏng các quá trình động học của graphene nanoribbon khuyết trong các điều kiện khác nhau. Tất cả các cài đặt huấn luyện mô hình cũng như trực quan hóa kết quả đều được chạy trên Google Colab bởi ngôn ngữ Python.

3. Kết quả và thảo luận

Tệp dữ liệu OUTCAR được đọc và phân chia tỉ lệ phần trăm dữ liệu cho tập huấn luyện và tập kiểm tra, bao gồm các tỉ lệ 60%, 70%, 80% và 90% cho tập huấn luyện tương ứng với 40%, 30%, 20% và 10% trên tập kiểm tra. Sau đó tiến hành huấn luyện cho từng mô hình trong DeePMD-kit, phân tích dữ liệu và trực quan hóa kết quả bằng gói Ovito. Để có thể đánh giá quá trình huấn luyện của các mô hình, đầu tiên chúng tôi vẽ biểu đồ thể hiện đường cong học, thông qua sự thay đổi của hàm mất mát (Loss) theo số bước (Step) được thể hiện ở Hình 2.

Hình 2a tỷ lệ 60% huấn luyện, 40% kiểm tra có $rmse_trn$ (cam) và $rmse_val$ (xanh dương) giảm đều, nhưng $rmse_val$ có dao động nhiều hơn, $rmse_e_trn$ (đỏ) giảm mạnh sau **100 bước**, nhưng $rmse_e_val$ (xanh lá) vẫn dao động nhiều, mô hình có dấu hiệu **quá khớp (overfitting)** khi $rmse_trn$ giảm ổn định nhưng $rmse_val$ biến động. Hình 2b tỷ lệ 70% huấn luyện, 30% kiểm tra, có thể thấy xu hướng tương tự 60%-40%, nhưng $rmse_val$ có dao động ít hơn. cho thấy tập kiểm tra lớn hơn giúp giảm biến động đáng kể, $rmse_e_trn$ (đỏ) giảm mạnh, nhưng $rmse_e_val$ (xanh lá) vẫn dao động mà không giảm, mặt khác $rmse_f_val$ và $rmse_f_trn$ (lục) vẫn duy trì ổn định. Đối với tỉ lệ 80% huấn luyện, 20% kiểm tra ở Hình 2c, có thể nhận thấy $rmse_trn$ và $rmse_val$ giảm gần nhau hơn, **ít bị overfitting hơn so với 60%-40%, 70%-30%, ngoài ra** $rmse_e_val$ giảm đều hơn, ít dao động, do đó có thể thấy mô hình 80%-20% cho kết quả **tổng quát hóa tốt nhất trong bốn mô hình khảo sát. Mặt khác** mô hình 80%-20% có thể dự đoán lực khá chính xác do $rmse_f_trn$ và $rmse_f_val$ gần như trùng nhau. Cuối cùng với mô hình có tỉ lệ 90% huấn luyện, 10% kiểm tra (Hình 2d), $rmse_trn$ và $rmse_val$ giảm tương tự nhau, nhưng $rmse_val$ có một số dao động mạnh cuối quá trình huấn luyện, $rmse_e_trn$ và $rmse_e_val$ giảm chậm hơn so với 80%-20%, $rmse_f_trn$ và $rmse_f_val$ dao động mạnh hơn so với 80%-20%. Có thể thấy rằng việc thiếu dữ liệu kiểm tra khiến mô hình 90%-10% khó đánh giá đầy đủ và tổng quát hóa. Với bốn mô hình khảo sát có thể kết luận **80%-20% là tỷ lệ tốt nhất** để huấn luyện mô hình DeePMD-kit trong trường hợp khảo

sát. Các kết quả so sánh về hiệu suất huấn luyện, hiệu suất kiểm tra và độ ổn định của các tỷ lệ phân chia được tóm tắt trong Bảng 1.



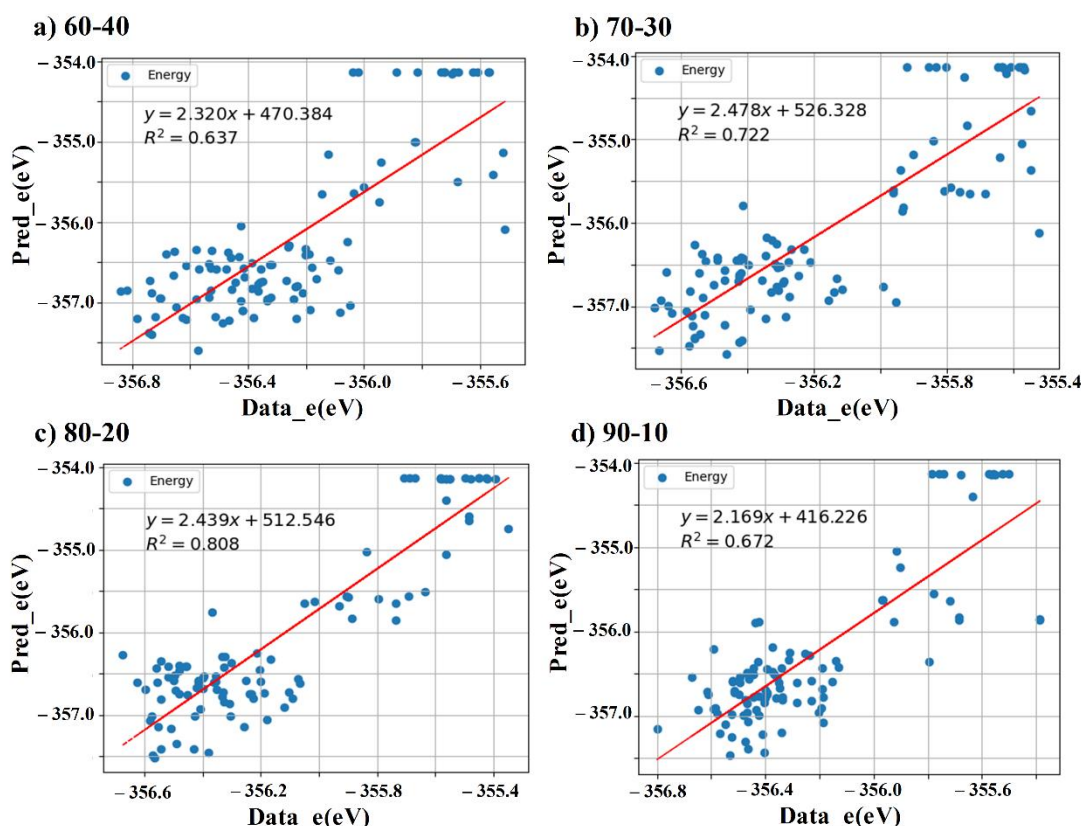
Hình 2. Biểu đồ thể hiện đường cong học của các mô hình huấn luyện, tỉ lệ phân chia tập huấn luyện-tập kiểm tra. a) 60%-40%, b) 70%-30%, c) 80%-20% và d) 90%-10%.

Trong đó, rmse_val (xanh dương): mất mát RMSE (root mean squared error) trên tập kiểm tra, rmse_trn (cam): mất mát RMSE trên tập huấn luyện, rmse_e_val (xanh lá cây): mất mát năng lượng trên tập kiểm tra, rmse_e_trn (đỏ): mất mát năng lượng trên tập huấn luyện, rmse_f_val (tím): mất mát lực trên tập kiểm tra, rmse_f_trn (nâu): mất mát lực trên tập huấn luyện.

Bảng 1. So sánh hiệu suất huấn luyện của các mô hình

Tỷ lệ	Hiệu suất huấn luyện	Hiệu suất kiểm tra	Độ ổn định
60-40	Nhanh nhưng có dấu hiệu overfitting	Kiểm tra dao động nhiều	Trung bình
70-30	Giảm overfitting, vẫn còn dao động	Kiểm tra ít dao động hơn	Tốt hơn 60-40
80-20	Huấn luyện tốt, không bị overfitting	Kiểm tra ổn định	Tốt nhất
90-10	Huấn luyện nhanh nhưng thiếu dữ liệu kiểm tra	Kiểm tra dao động mạnh	Không ổn định

Để có thể đánh giá các mô hình đầy đủ hơn, chúng tôi tiến hành vẽ đồ thị hiển thị mối quan hệ giữa năng lượng dự đoán (Pred_e) và năng lượng thực tế (Data_e) với bốn tỷ lệ phân chia dữ liệu huấn luyện và kiểm tra được khảo sát trong Hình 3.



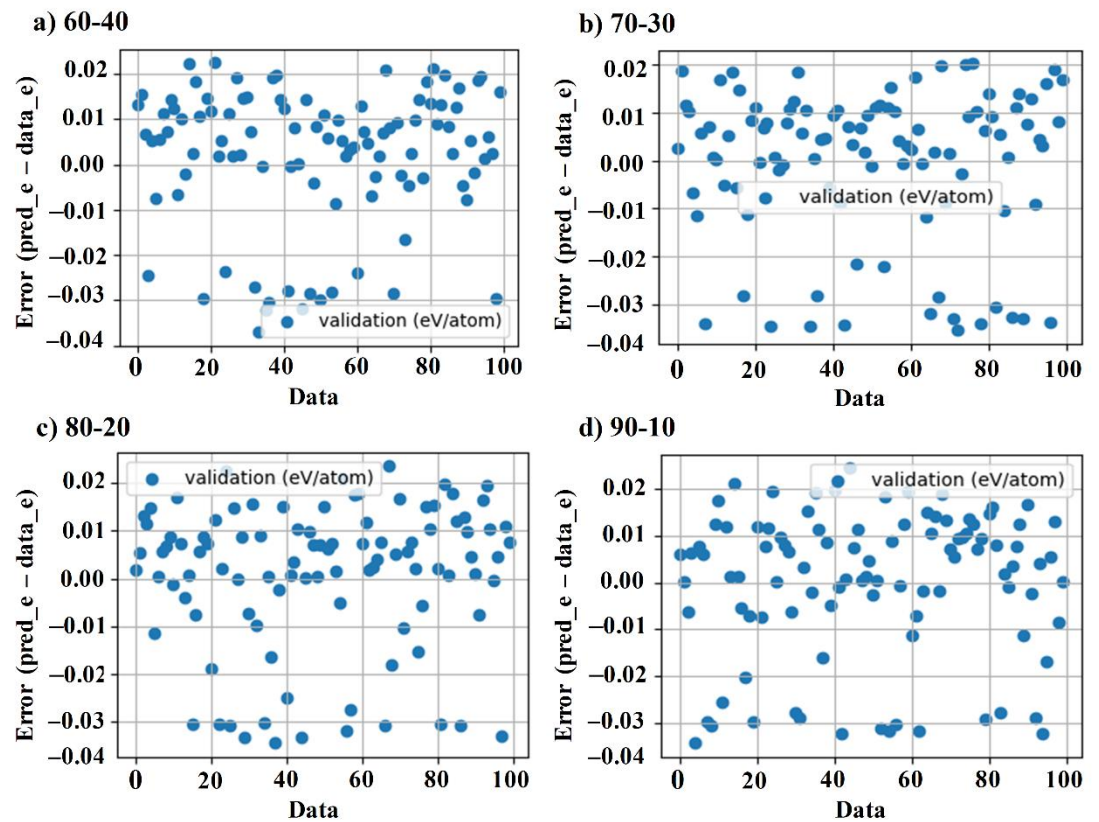
Hình 3. Mối quan hệ giữa năng lượng dự đoán (Pred_e) và năng lượng thực tế (Data_e) với bốn tỷ lệ phân chia dữ liệu huấn luyện và kiểm tra. a) 60%-40%, b) 70%-30%, c) 80%-20%, d) 90%-10%.

Hình 3 với trục ngang (Data_e) là giá trị năng lượng thực tế trong tệp OUTCAR (tính từ AIMD/VASP), trục đứng (Pred_e) tương ứng giá trị năng lượng dự đoán của mô hình MLIP. Các điểm dữ liệu được tô màu xanh minh họa năng lượng thực tế so với năng lượng dự đoán, đường hồi quy tuyến tính (màu đỏ) dạng $y = ax + b$. Các mô hình được đánh giá thông qua hệ số xác định R^2 , với R^2 càng gần về một thì mô hình càng chính xác. Với bốn mô hình khảo sát, mô hình 80%-20% có các dữ liệu bám sát đường hồi quy, cùng với giá trị $R^2 = 0.808$ là lớn nhất so với các mô hình còn lại, cho thấy đây là mô hình tốt nhất đạt độ chính xác cao hơn so với các mô hình khảo sát, điều này phù hợp với các kết quả thể hiện đường cong học có được ở Hình 2. Ở ba mô hình còn lại với các giá trị R^2 là 0.637, 0.722, 0.672 lần lượt cho các tỷ lệ phân chia 60%-40%, 70%-30% và 90%-10%, có thể thấy các dữ liệu phân tán rộng quanh đường hồi quy nên các mô hình này kém chính xác hơn so với mô hình 80%-20%. Mô hình 90%-10% do tập kiểm tra có tỉ lệ nhỏ nên mô hình có thể bị quá khớp (overfitting) với tập huấn luyện, dẫn đến giảm độ chính xác. Các thống kê kết quả và hiệu suất của từng mô hình được thống kê ở Bảng 2.

Bảng 2. Thống kê so sánh hệ số xác định và hiệu suất các mô hình

Tỷ lệ tập dữ liệu	Hệ số xác định R^2	Hiệu suất
60-40	0.637	Trung bình, dữ liệu phân tán nhiều.
70-30	0.722	Cải thiện, nhưng vẫn có sai số.
80-20	0.808	Tốt nhất, dự đoán chính xác cao.
90-10	0.672	Quá khớp (overfitting), độ chính xác giảm.

Để có cái nhìn trực quan hơn về các mô hình, chúng tôi đánh giá thêm độ lệch của giá trị năng lượng dự đoán và năng lượng thực tế được thể hiện trong Hình 4.



Hình 4. Độ lệch giữa năng lượng dự đoán (Pred_e) và năng lượng thực tế (Data_e) với bốn tỷ lệ phân chia dữ liệu huấn luyện và kiểm tra. a) 60%-40%, b) 70%-30%, c) 80%-20%, d) 90%-10%.

Có thể nhận thấy, mô hình 60%-40% có các điểm dữ liệu về độ lệch năng lượng (Error) phân tán rộng từ khoảng -0.04 eV/atom đến 0.02 eV/atom. Sự phân tán này cho thấy độ chính xác của mô hình khá thấp, với nhiều điểm có sai số lớn. Với mô hình 70%-30% các điểm dữ

liệu có độ lệch năng lượng dao động từ khoảng -0.03 eV/atom đến 0.02 eV/atom. Tuy có một số điểm lệch, nhưng sự phân tán này ít hơn so với mô hình 60%-40%. Còn với mô hình 80%-20% ta thấy độ lệch năng lượng chủ yếu trong khoảng từ -0.03 eV/atom đến 0.01 eV/atom, các điểm dữ liệu bám khá chặt vào khoảng sai số nhỏ và gần về 0.00 eV/atom cho thấy mô hình dự đoán khá chính xác. Cuối cùng mô hình 90%-10% các điểm dữ liệu của mô hình này có độ lệch năng lượng dao động từ khoảng -0.03 eV/atom đến 0.02 eV/atom, với một số điểm có sai số lớn, Mô hình này có thể bị quá khớp (overfitting) do tỷ lệ dữ liệu kiểm tra quá nhỏ (10%). Mặc dù các điểm chủ yếu nằm trong phạm vi sai số nhỏ, nhưng độ lệch năng lượng có sự phân tán lớn, cho thấy mô hình có thể dự đoán không ổn định.

4. Kết luận

Trong nghiên cứu này, chúng tôi đã khảo sát hiệu suất của bốn mô hình huấn luyện MLIP với các tỷ lệ phân chia dữ liệu tập huấn luyện-tập kiểm tra khác nhau lần lượt là 60%-40%, 70%-30%, 80%-20% và 90%-10%, để mô phỏng động lực học phân tử của graphene nanoribbon khuyết. Các mô hình được huấn luyện sử dụng DeePMD-kit, với các thông số cấu trúc và năng lượng từ mô phỏng AIMD/VASP. Kết quả thu được đã cung cấp những đánh giá quan trọng về khả năng dự đoán năng lượng và lực, cũng như sự ổn định của các mô hình.

Các kết quả cho thấy mô hình với tỷ lệ phân chia 80%-20% tập huấn luyện-tập kiểm tra có hiệu suất tốt nhất, với độ chính xác cao và sự giảm dần ổn định của hàm mất mát trên cả tập huấn luyện và kiểm tra. Cụ thể, mô hình này đạt giá trị $R^2 = 0.808$, phản ánh khả năng dự đoán năng lượng chính xác và ít bị phân tán sai số. Mô hình này cũng cho thấy độ lệch năng lượng thấp và ổn định, với các điểm dữ liệu gần đường hồi quy. Ngoài ra các mô hình 60%-40% và 70%-30% có sự phân tán sai số lớn hơn, đặc biệt là ở mô hình 60%-40%, với độ lệch năng lượng dao động mạnh, cho thấy độ chính xác thấp. Mô hình 90%-10%, mặc dù huấn luyện nhanh, nhưng thiếu dữ liệu kiểm tra khiến mô hình không thể tổng quát hóa tốt và dẫn đến độ lệch năng lượng không ổn định.

Các kết quả nghiên cứu cung cấp một cái nhìn sâu sắc về hiệu suất của các mô hình MLIP trong việc dự đoán năng lượng và lực của vật liệu nanoribbon, mở ra hướng nghiên cứu tiếp theo về tối ưu hóa các mô hình với các thể tương tác mới cho các hệ vật liệu khác nhau.

Lời cảm ơn: Nghiên cứu được tài trợ bởi Trường Đại học Nam Cần Thơ, mã số: C24.30. Nghiên cứu được thực hiện bởi hệ thống máy tính hiệu năng cao của Lab MOMA, tòa nhà công nghệ cao ATL, Trường Đại học Cần Thơ.

Tài liệu tham khảo

- Behler, J. (2016). Perspective: Machine learning potentials for atomistic simulations. *The Journal of Chemical Physics*, 145(17), 170901. <https://doi.org/10.1063/1.4966192>
- Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O., & Walsh, A. (2018). Machine learning for molecular and materials science. *Nature*, 559(7715), 547–555. <https://doi.org/10.1038/s41586-018-0337-2>
- Hafner, J. (2008). Ab-initio simulations of materials using VASP: Density-functional theory and beyond. *Journal of computational chemistry*, 29(13), 2044–2078. <https://doi.org/10.1002/jcc.21057>
- Jinnouchi, R., Karsai, F., & Kresse, G. (2019). On-the-fly machine learning force field generation: Application to melting points. *Physical Review B*, 100(1), 014105. <https://doi.org/10.1103/PhysRevB.100.014105>

- Nguyen, T. T., Pham, T. B. T., Nguyen, D. K., Le, N. T., & Vo, K. D. (2025). Thermoelectric properties of penta-InPs: A first-principles and machine learning study. *Journal of Applied Physics*, 137(8), 084302. <https://doi.org/10.1063/5.0251741>
- Nguyen, T. T., Vo, T. P., & Ahuja, R. (2018). Tuning electronic transport properties of zigzag graphene nanoribbons with silicon doping and phosphorus passivation. *AIP Advances*, 8(8), 085123. <https://doi.org/10.1063/1.5042071>
- Schmidt, J., Marques, M. R. G., Botti, S., & Marques, M. A. L. (2019). Recent advances and applications of machine learning in solid-state materials science. *npj Computational Materials*, 5, 83. <https://doi.org/10.1038/s41524-019-0221-0>
- Shapeev, A. V. (2016). Moment tensor potentials: A class of systematically improvable interatomic potentials. *Multiscale Modeling & Simulation*, 14(3), 1153–1173. <https://doi.org/10.1137/15M1054183>
- Smidstrup, S., Markussen, T., Vancraeyveld, P., Wellendorff, J., Schneider, J., Gunst, T., ... & Stokbro, K. (2019). QuantumATK: an integrated platform of electronic and atomic-scale modelling tools. *Journal of Physics: Condensed Matter*, 32(1), 015901. <https://doi.org/10.1088/1361-648X/ab4007>
- Wang, H., Zhang, L., & Han, J. (2018). DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Computer Physics Communications*, 228, 178-184. <https://doi.org/10.1016/j.cpc.2018.03.016>
- Yazyev, O. V., & Louie, S. G. (2010). Electronic transport in polycrystalline graphene. *Nature Materials*, 9(10), 806–809. <https://doi.org/10.1038/nmat2830>
- Zeng, J., Giese, T. J., Zhang, D., Wang, H., & York, D. M. (2025). DeePMD-GNN: A DeePMD-kit Plugin for External Graph Neural Network Potentials. *Journal of Chemical Information and Modeling*, 65(7), 3154-3160. <https://doi.org/10.1021/acs.jcim.4c02441>
- Zhang, Y., Li, H., & Wang, J. (2022). Various defects in graphene: a review. *RSC Advances*, 12(24), 15644–15654. <https://doi.org/10.1039/D2RA01436J>
- Zhao, Y., Liu, Z., & Wang, Y. (2022). Effect of Vacancy Defects on the Vibration Frequency of Graphene Nanoribbons. *Nanomaterials*, 12(5), 764. <https://doi.org/10.3390/nano12050764>
- Zuo, Y., Chen, C., Li, X., Deng, Z., Chen, Y., Behler, J., ... & Ong, S. P. (2020). Performance and cost assessment of machine learning interatomic potentials. *Journal of Physical Chemistry A*, 124(4), 731-745. <https://doi.org/10.1021/acs.jpca.9b08723>