

Research Paper

Prediction of Ion Adsorption Free Energy on CNT/GO Using a Hybrid Molecular Dynamic and Random Forest

Hoda Ghavaminia *

Physics Department, Materials and Energy Research Center, Dez. Co., Islamic Azad University, Dezful, Iran

Received: 2025/09/09

Accepted: 2025/11/23

Use your device to scan and read the
article online

**Keywords:**

Ion adsorption, Graphene oxide, Molecular dynamics, Machine learning, Free energy, Heavy metals.

Abstract

Introduction: Removing heavy metal ions from water is a major environmental challenge. This study aimed to predict the adsorption free energy of these ions onto a carbon nanotube/graphene oxide (CNT/GO) nanocomposite.

Methods: A hybrid methodology integrating molecular dynamics (MD) simulations with machine learning (ML) models was developed and applied. The free energy of adsorption (ΔG_{ads}) was calculated using umbrella sampling and weighted histogram analysis. Various ML models, including Random Forest, Kernel Ridge Regression, and Multilayer Perceptron, were developed and evaluated.

Findings: The findings revealed a clear difference in the ions' thermodynamic affinity. Lead ions had the highest adsorption tendency, while cadmium had the lowest. Microscopic analysis confirmed strong chemical bonding for lead and copper with surface functional groups, whereas cadmium only showed weak physical interaction. The Random Forest model demonstrated the best predictive performance. Analysis of feature importance identified ionic charge as the most critical parameter.

Conclusion: This research provides an effective hybrid framework for predicting adsorption behavior. The developed model successfully captured complex structure-property relationships, enabling the targeted design of selective adsorbents. This approach can significantly reduce the time and cost of developing new water purification materials.

* **Corresponding author:** Hoda Ghavaminia

Email: h.ghavaminia@iau.ac.ir

DOI: <https://doi.org/10.71905/jnm.2025.1217479>

Extended Abstract

Introduction:

The removal of toxic heavy metal ions from aqueous environments is a critical challenge in environmental science. Carbon-based nanocomposites, particularly carbon nanotube/graphene oxide (CNT/GO) structures, have emerged as promising adsorbents due to their exceptional surface properties and functionalization capabilities. However, understanding the atomic-scale mechanisms and thermodynamics of ion adsorption remains computationally challenging. This study develops a hybrid molecular dynamics-machine learning framework to quantitatively predict adsorption behavior and identify key factors governing ion selectivity on CNT/GO surfaces.

Findings and Discussion

In this study, a hybrid molecular dynamics (MD) and machine learning (ML) approach was employed to quantitatively investigate the adsorption behavior of Pb^{2+} , Cu^{2+} , and Cd^{2+} ions on a carbon nanotube/graphene oxide (CNT/GO) nanocomposite. The adsorption free energy (ΔG_{ads}) was precisely calculated using the umbrella sampling technique combined with the Weighted Histogram Analysis Method (WHAM). The results revealed a distinct thermodynamic affinity trend: Pb^{2+} (-12.5 kcal/mol) > Cu^{2+} (-9.8 kcal/mol) > Cd^{2+} (-8.3 kcal/mol), indicating spontaneous adsorption for all ions with a strong preference for lead. This trend aligns remarkably well with both experimental adsorption capacity data and previous theoretical DFT calculations, validating our simulation protocol. Radial distribution function (RDF) analysis and spatial density profiles provided profound insights into the molecular-scale adsorption mechanisms. To gain deeper mechanistic insights beyond this validation, radial distribution function (RDF) analysis and spatial density profiles were employed. These analyses provided profound insights into the molecular-scale adsorption mechanisms. RDF peaks between 2.0–2.5 Å for Pb^{2+} and Cu^{2+} ions and oxygen atoms of surface functional groups (-OH, -COOH) confirmed the formation of strong, directional coordination bonds, characteristic of chemisorption. In contrast, the RDF for Cd^{2+} showed a broad, weak peak at a larger distance (~3.5 Å), signifying predominantly physical, electrostatic interactions (physisorption) with a more diffuse hydration shell. To predict ΔG_{ads}

efficiently, three machine learning models—Random Forest (RF), Kernel Ridge Regression (KRR), and Multilayer Perceptron (MLP)—were developed. The Random Forest model demonstrated superior performance, achieving an R^2 of 0.92 and an RMSE of 0.45 kcal/mol, underscoring its capability to capture the complex, non-linear relationships between ion characteristics and adsorption energy. Feature importance analysis, utilizing both Gini importance and SHAP values, identified **ionic charge** as the most critical predictor, followed by **ionic radius** and **Lennard-Jones parameters** (ϵ and σ). This highlights that adsorption is primarily governed by the strength of electrostatic interactions and the ion's ability to approach the surface closely, which is consistent with the physical chemistry of ion adsorption.

Conclusion

This study establishes an effective MD-ML framework for predicting and analyzing heavy metal adsorption on CNT/GO nanocomposites. The results demonstrate clear correlation between ion characteristics and adsorption energy, with Pb^{2+} showing strongest affinity due to its high charge density and coordination capability. The Random Forest model proved particularly effective in capturing complex structure-property relationships, providing valuable insights for rational design of selective adsorbents. This approach offers a powerful tool for accelerating development of advanced water purification materials while reducing reliance on experimental trial-and-error methods.

Ethical Considerations compliance with ethical guidelines

All data were generated through simulations and analysis of publicly available or computationally derived information. Therefore, no ethical approval was required.

Funding

No funding.

Authors' contributions

The author confirms sole responsibility

Conflicts of interest

The authors declared no conflict of interest.

مقاله پژوهشی

پیش‌بینی انرژی آزاد جذب یون‌ها بر روی CNT/GO با استفاده از ترکیب دینامیک مولکولی و مدل جنگل تصادفی

هدی قوامی نیا *

استادیار، گروه فیزیک، مرکز تحقیقات مواد و انرژی، واحد دزفول، دانشگاه آزاد اسلامی، دزفول، ایران

چکیده

مقدمه: جذب یون‌های فلزات سنگین از محیط‌های آبی یک چالش زیست‌محیطی مهم محسوب می‌شود. این مطالعه با هدف پیش‌بینی انرژی آزاد جذب این یون‌ها بر روی نانوکامپوزیت نانولوله کربنی/اکسید گرافن (CNT/GO) انجام شد.

روش: در این تحقیق از یک رویکرد ترکیبی شامل شبیه‌سازی‌های دینامیک مولکولی و مدل‌های یادگیری ماشین استفاده گردید. انرژی آزاد جذب با استفاده از روش نمونه‌برداری چتری و تحلیل هیستوگرام وزندار محاسبه شد. مدل‌های مختلف یادگیری ماشین شامل جنگل تصادفی، رگرسیون ریبج هسته‌ای و پرسپترون چندلایه توسعه و ارزیابی شدند.

یافته‌ها: یافته‌ها تمایز واضحی در تمایل ترمودینامیکی یون‌ها نشان داد. یون سرب بیشترین و یون کادمیوم کمترین تمایل به جذب داشت. بررسی‌های میکروسکوپی تشکیل پیوندهای شیمیایی قوی برای سرب و مس را تأیید کرد، در حالی که کادمیوم تنها برهمکنش فیزیکی ضعیفی نشان داد. مدل جنگل تصادفی بهترین عملکرد پیش‌بینی را ارائه داد و تحلیل عوامل مؤثر، بار یونی را به عنوان مهم‌ترین پارامتر شناسایی کرد.

نتیجه‌گیری: این پژوهش یک چارچوب ترکیبی مؤثر برای پیش‌بینی رفتار جذب ارائه می‌دهد. مدل توسعه‌یافته توانایی بالایی در شناسایی روابط پیچیده بین ساختار و خواص نشان داد و زمینه را برای طراحی هدفمند جاذب‌های انتخابی فراهم کرد. این رویکرد می‌تواند زمان و هزینه‌های توسعه مواد جدید برای تصفیه آب را به‌طور چشمگیری کاهش دهد.

تاریخ دریافت: ۱۴۰۴/۰۶/۱۸

تاریخ پذیرش: ۱۴۰۴/۰۹/۰۲

از دستگاه خود برای اسکن و خواندن مقاله به صورت آنلاین استفاده کنید



واژه‌های کلیدی:

جذب یونی، اکسید گرافن، دینامیک مولکولی، یادگیری ماشین، انرژی آزاد، فلزات سنگین.

نویسنده مسئول: هدی قوامی نیا

پست الکترونیکی: h.ghavaminia@iaau.ac.ir

DOI: <https://doi.org/10.71905/jnm.2025.1217479>

مقدمه

آلودگی آب توسط یون‌های فلزات سنگین یکی از جدی‌ترین چالش‌های زیست‌محیطی در قرن حاضر است. فلزاتی مانند سرب (Pb^{2+})، کادمیوم (Cd^{2+}) و مس (Cu^{2+}) عمدتاً از فعالیت‌های صنعتی، معدن‌کاری، پساب‌های الکتروپلاستی و باتری‌ها وارد محیط زیست می‌شوند و به دلیل سمیت بالا، عدم تجزیه‌پذیری و تجمع‌پذیری در زنجیره غذایی، تهدیدی جدی برای سلامت انسان و اکوسیستم‌ها محسوب می‌شوند [۱]. بنابراین، توسعه روش‌های کارآمد و انتخابی برای حذف این یون‌ها از پساب‌های صنعتی و منابع آبی از اهمیت بالایی برخوردار است. در سال‌های اخیر، نانومواد کربنی به‌ویژه اکسید گرافن^۱ (GO) و کامپوزیت‌های مبتنی بر گرافن به‌عنوان جاذب‌های نوین و پرکاربرد در حذف یون‌های فلزات سنگین مورد توجه گسترده قرار گرفته‌اند [۲]. این مواد به دلیل سطح ویژه بالا، حضور گروه‌های عاملی قطبی مانند ($-OH$ ، $-COOH$ ، $-O-$)، و قابلیت اصلاح سطحی، تعاملات قوی با یون‌های فلزی از طریق مکانیسم‌های جذب سطحی، کمپلکس‌شدن و تبادل یونی ایجاد می‌کنند [۳]. با این‌حال، درک عمیق و کمی از مکانیسم‌های ترمودینامیکی و دینامیکی جذب، به‌ویژه در مقیاس اتمی، همچنان چالش برانگیز است. روش‌های تجربی مانند جذب سنتی و طیف‌سنجی، اگرچه اطلاعات ارزشمندی ارائه می‌دهند، اما در شناسایی جزئیات مولکولی مانند هندسه تعامل، انرژی آزاد جذب و نقش گروه‌های عاملی محدودیت دارند. در این زمینه، شبیه‌سازی دینامیک مولکولی^۲ (MD) به‌عنوان ابزاری قدرتمند برای مدل‌سازی رفتار سیستم‌های پیچیده در مقیاس زمانی و فضایی اتمی، امکان بررسی دقیق مکانیسم جذب را فراهم کرده است [۴]. به‌ویژه، استفاده از روش نمونه‌برداری چتری^۳ برای محاسبه انرژی آزاد جذب (ΔG_{ads}) امکان مقایسه کمی تمایل یون‌ها به جذب بر روی سطوح مختلف را فراهم می‌کند [۵]. با ظهور و توسعه فناوری‌های یادگیری ماشین^۴، پیوندی مؤثر بین شبیه‌سازی‌های پردازشی دقیق و مدل‌های پیش‌بین سریع و قابل اعتماد برقرار شده است. این مدل‌ها با کشف الگوهای نهفته در داده‌های حاصل از شبیه‌سازی‌ها، قادرند رفتار ترمودینامیکی سامانه‌ها را با صحت قابل توجهی پیش‌بینی نموده و هم‌زمان، پارامترهای کلیدی

تأثیرگذار بر مکانیسم جذب را به شکلی کمی و تفسیرپذیر آشکار سازند. امکان آ شبیه‌سازی‌های محاسباتی با مدل‌های پیش‌بینی سریع و دقیق فراهم شده است. مدل‌های یادگیری ماشین می‌توانند با یادگیری از داده‌های شبیه‌سازی، رفتار ترمودینامیکی سیستم‌ها را پیش‌بینی کنند و ویژگی‌های کلیدی مؤثر بر جذب را شناسایی نمایند [۶]. این رویکرد ترکیبی دینامیک مولکولی-یادگیری ماشین نه تنها دقت پیش‌بینی را افزایش می‌دهد، بلکه چرخه طراحی مواد جاذب جدید را به‌طور چشمگیری تسریع می‌کند. در این مطالعه، از ترکیب شبیه‌سازی دینامیک مولکولی و یادگیری ماشین برای بررسی جذب یون‌های Pb^{2+} ، Cu^{2+} و Cd^{2+} بر روی سطوح کامپوزیت نانولوله کربنی/اکسید گرافن^۵ (CNT/GO) استفاده شده است. ابتدا با استفاده از روش نمونه‌گیری چتری، انرژی آزاد جذب هر یون محاسبه شد. سپس، مدل‌های یادگیری ماشین همچون مدل جنگل تصادفی^۶، رگرسیون هسته ریب^۷ و پرسپترون چند لایه^۸ برای پیش‌بینی ΔG_{ads} بر اساس ویژگی‌های فیزیکی و شیمیایی آموزش داده شدند. در نهایت، اهمیت ویژگی‌ها با استفاده از روش شاپ^۹ و ضریب جنی^{۱۰} تحلیل شد تا عوامل کنترل‌کننده اصلی جذب شناسایی شوند. بدین ترتیب، چارچوبی جامع برای درک و پیش‌بینی رفتار جذب یون‌های فلزی بر پایه ویژگی‌های اتمی و سطحی فراهم آمد. ساختار این مقاله به شرح زیر است: در بخش ۲، روش‌شناسی شامل جزئیات شبیه‌سازی دینامیک مولکولی، محاسبه انرژی آزاد و مدل‌سازی یادگیری ماشین ارائه شده است. بخش ۳ به نتایج و بحث اختصاص دارد که شامل تحلیل ترمودینامیکی، اعتبارسنجی مدل و مکانیسم جذب است. در بخش ۴، نتیجه‌گیری و چشم‌انداز آینده ارائه شده و کاربردهای عملی یافته‌ها برای طراحی جاذب‌های هوشمند برجسته شده است.

روش‌ها

۱- آماده‌سازی سیستم‌های شبیه‌سازی

سیستم‌های شبیه‌سازی شامل لایه‌های تک‌مولکولی گرافن و اکسید گرافن بودند که در محیط آبی حاوی یون‌های فلزات سنگین دوظرفیتی Pb^{2+} ، Cu^{2+} و Cd^{2+} قرار داشتند. برای مدل‌سازی ساختار گرافن، ابعاد جانبی 6×6 نانومتر مربع در نظر گرفته شد تا اثرات مرزی به حداقل برسد [۷]. اکسید گرافن با

- 1 Graphene Oxide
- 2 Molecular Dynamics
- 3 Umbrella Sampling
- 4 Machine Learning
- 5 carbon nanotube/graphene oxide (CNT/GO)
- 6 Random Forest

- 7 KRR- Kernel Ridge Regression
- 8 MLP-Multi Layer Perceptron
- 9 SHAP-SHapley Additive exPlanations
- 10 Gini Importance

جزئی اتم‌ها، r_{ij} ، فاصله بین دو اتم، ϵ_0 ، ثابت دی‌الکتریک خلأ می‌باشد. برای محاسبه تعاملات الکترواستاتیکی دوربرد، از روش^۵ PPPM با دقت عددی 10^{-10} کیلوکالری بر مول استفاده شد. همچنین، با توجه به هندسه لایه‌ای سیستم، اصلاحات لایه‌ای با دستور `kspace\_modify slab 3.0` اعمال گردید تا اثرات نامحدود در جهت عمودی به‌درستی مدنظر قرار گیرد [۱۵]. زمان‌گام شبیه‌سازی ۱ فمتوثانیه در نظر گرفته شد و از الگوریتم شیک_جبرای قیود طول پیوند در مولکول‌های آب استفاده شد تا پایداری عددی شبیه‌سازی افزایش یابد [۱۶].

جدول ۱-الف: پارامترهای میدان نیروی جونگ-چیام برای

یون‌های فلزی			
ϵ (kcal/mol)	σ (Å)	بار (e)	یون
۰/۰۲۷	۳/۶۸	+۲/۰	Pb ²⁺
۰/۰۵۰	۲/۴۶	+۱/۸۰	Cu ²⁺
۰/۰۴۰	۲/۵۴	+۱/۵۰	Cd ²⁺

جدول ۱-ب: پارامترهای میدان نیرو و بارهای جزئی برای

اتم‌های سطح CNT/GO				
ϵ (kcal/mol)	σ (Å)	بار (e)	نوع اتم	گروه عاملی
۰/۰۷۰	۳/۵۵	۰/۰	C	گرافن
۰/۱۷۰	۳/۱۲	-۰/۷۰	O	هیدروکسیل (-OH)
۰/۰۰	۰/۰۰	+۰/۴۳	H	هیدروکسیل (-OH)
۰/۱۰۵	۳/۷۵	+۰/۵۸	C (کربونیل)	کربوکسیل -COOH
۰/۲۱۰	۲/۹۶	-۰/۵۰	O (کربونیل)	کربوکسیل -COOH
۰/۱۷۰	۳/۰۰	-۰/۶۵	O (هیدروکسیل)	کربوکسیل -COOH
۰/۰۰	۰/۰۰	+۰/۴۵	H (هیدروکسیل)	کربوکسیل -COOH
۰/۱۷۰	۳/۰۰	-۰/۴۰	O	اپوکسی -O-
۰/۰۷۰	۳/۵۵	+۰/۲۰	C	اپوکسی -O-

۳. تعادل‌سازی سیستم

برای رسیدن به حالت تعادل ترمودینامیکی، با استفاده از ترکیب روش‌های نزول گرادیان و گرادیان مختلط مینیمم‌سازی انرژی انجام شد تا سیستم به کمینه موضعی انرژی پتانسیل برسد [۱۷]. سیستم از دمای صفر کلون به دمای هدف ۲۹۸ کلون در مدت ۲۰۰ پیکوثانیه در مجموعه تعداد ذرات، حجم و دمای ثابت گرم

توزیع تصادفی گروه‌های عاملی شامل هیدروکسیل، اپوکسی و کربوکسیل روی صفحه گرافن ایجاد شد. نسبت اکسیژن به کربن O/C در دو سطح ۰/۲ و ۰/۳۵ تنظیم گردید تا اثر گرادیان قطبیت سطح بر جذب یون‌ها بررسی شود [۸]. سیستم‌ها با یون‌های فلزی با غلظت اولیه ۰/۵ مولار اشباع شدند و برای حفظ خنثایی بار الکتریکی، یون‌های متقابل Cl^- به‌صورت تصادفی در حلال توزیع شدند. شرایط مرزی دوره‌ای در جهت‌های x و y اعمال شد، در حالی که در جهت z، یک لایه خلأ به ضخامت حداقل ۶ نانومتر در نظر گرفته شد تا از تعاملات میان تصویری^۱ در جهت عمود بر صفحه جلوگیری شود [۹]. این ضخامت خلأ به‌اندازه‌ای کافی بود که از تأثیرات غیرفیزیکی ناشی از تصاویر دور در شبیه‌سازی‌های دو بعدی جلوگیری کند.

۲. میدان‌های نیرو و پارامترهای شبیه‌سازی

تمامی شبیه‌سازی‌های دینامیک مولکولی با استفاده از کد محاسباتی لمپس^۲ انجام شد [۱۰]. مدل TIP3P برای شبیه‌سازی مولکول‌های آب به‌کار گرفته شد، که دقت مناسبی در توصیف خواص دینامیکی و ساختاری آب در دمای اتاق دارد [۱۱]. برای یون‌های فلزی Pb^{2+} ، Cu^{2+} و Cd^{2+} ، پارامترهای میدان نیروی جونگ-چیام^۳ استفاده شد که به‌طور گسترده در مطالعات حلال‌سازی و جذب یونی اعتبارسنجی شده است [۱۲]. ساختار گرافن و اکسید گرافن با استفاده از میدان نیروی OPLS-AA مدل‌سازی شد [۱۳]. پارامترهای نهایی این یون‌ها در جدول ۱-الف ارائه شده است. بارهای جزئی گروه‌های عاملی اکسید گرافن از داده‌های محاسبه‌شده توسط روش‌های تئوری تابعی چگالی با تابع پایه B3LYP/6-31G و روش CHelpG استخراج گردید [۱۴]. بارهای محاسبه‌شده با مقادیر گزارش‌شده در مطالعات مشابه مطابقت خوبی نشان داد [۱۴]. این رویکرد، بازتولید واقع‌گرایانه‌ای از برهمکنش‌های الکترواستاتیکی مؤثر در جذب یون‌ها را ممکن ساخته است. پارامترهای کامل سطح در جدول ۱-ب آمده است.

تعاملات غیرکووالانسی با استفاده از پتانسیل لِنارد-جونز^۴ تعریف شد:

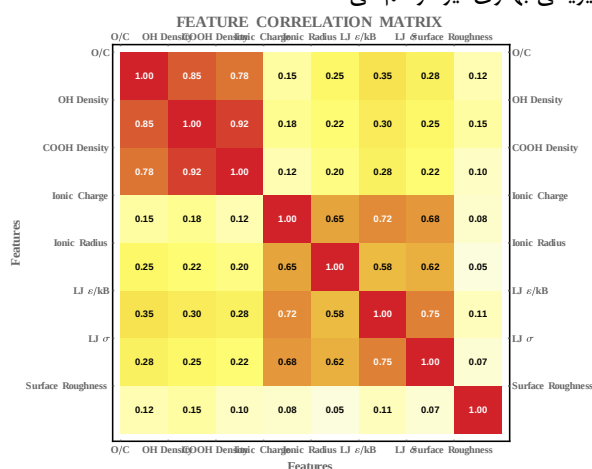
$$U_{ij}(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad (1)$$

که در آن ϵ_{ij} عمق چاه انرژی تعامل بین ذرات (اُ ج)، σ_{ij} فاصله‌ای که در آن انرژی تعامل صفر می‌شود، $q_i q_j$ بارهای

5 Particle-Particle Particle-Mesh
6 SHAKE

1 image-image interactions
2 LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator)
3 Joung-Cheatham
4 Lennard-Jones

نماینده از هر گروه ویژگی‌های همبسته در مدل نهایی حفظ گردید. برای افزایش تعمیم‌پذیری مدل از سه راهکار اصلی بهره‌گیری از ساختار ذاتی مقاوم جنگل تصادفی در برابر داده‌های همبسته، اعمال رگولایزاسیون L2 در شبکه عصبی مصنوعی، اجرای اعتبارسنجی متقابل ۵-تایی در فرآیند آموزش و بهینه‌سازی استفاده شد. همچنین کارایی روش کاهش ابعاد PCA مورد آزمایش قرار گرفت که علی‌رغم کاهش مؤثر ابعاد، بهبود محسوسی در دقت مدل جنگل تصادفی ایجاد نکرد. در مقابل این موضوع نشان می‌دهد رویکرد انتخاب ویژگی مبتنی بر همبستگی، علاوه بر حفظ قدرت پیش‌بینی، قابلیت تفسیرپذیری فیزیکی بهتری نیز فراهم می‌کند.



شکل ۱- ماتریس همبستگی ویژگی‌های مورد استفاده در مدل‌های یادگیری ماشین. مقادیر نشان‌دهنده ضرایب همبستگی پیرسون هستند. همبستگی بالا |۳| بزرگتر از ۰/۷ بین گروه‌های عاملی سطحی مشاهده می‌شود

۶. استخراج ویژگی‌های یادگیری ماشین

مجموعه‌ای از ویژگی‌های فیزیکی، شیمیایی و ساختاری شامل ویژگی‌های سطحی همچون نسبت O/C، چگالی گروه‌های عاملی (-COOH، -OH، -O) بر حسب nm^{-2} زبری سطح (انحراف معیار موقعیت اتم‌های کربن)، بار مؤثر سطح و ویژگی‌های یونی همچون شعاع یونی، بار مؤثر، قطبش‌پذیری، پارامترهای لِنارد-جونز (ϵ ، σ) و ویژگی‌های تعاملی همچون تابع توزیع شعاعی بین یون و اتم‌های اکسیژن سطح یا اکسیژن آب، عدد کوآرڈیناسیون CN^۵ در پوسته اول هیدراتاسیون، پروفایل چگالی یون‌ها در جهت z با استفاده از اسکرپت‌های سفارشی پایتون^۶ و وی ام دی^۷ استخراج شدند [۲۴] تا به‌عنوان ورودی مدل‌های یادگیری ماشین استفاده شوند. داده‌های استخراج‌شده

شد. کنترل دما با استفاده از ترموستات برندنسن^۱ انجام شد [۱۸]. پس از گرمایش، شبیه‌سازی ۲ نانوثانیه در مجموعه NVT ادامه یافت تا پارامترهای ترمودینامیکی (انرژی کل، دما و فشار) همگرا شوند. همگرایی سیستم با نظارت بر انرژی پتانسیل کل، انحراف معیار مربعات ریشه موقعیت اتم‌ها نسبت به حالت مرجع سطح گرافن و چگالی سیستم تأیید شد.

۴. محاسبه انرژی آزاد جذب ΔG_{ads}

برای محاسبه انرژی آزاد جذب یون‌ها بر روی سطح، از روش نمونه‌برداری چتری استفاده شد [۱۹]. مختصه واکنش به‌عنوان فاصله عمودی میان مرکز جرم یون و سطح گرافن -اکسید گرافن تعریف شد:

$$\xi = Z_{\text{ion}} - Z_{\text{surface}} \quad (2)$$

یک سری پنجره‌های همپوشان با گام ۱-۲ آنگستروم در بازه (30 تا 2 آنگستروم ξ) ایجاد شد. در هر پنجره، یک پتانسیل هارمونیک برای محدود کردن حرکت یون حول موقعیت مرجع ξ_0 اعمال شد:

$$U_{\text{bias}}(\xi) = \frac{1}{2}k(\xi - \xi_0)^2 \quad (3)$$

که در آن $k = 1000 \text{ kcal/mol} \cdot \text{\AA}^2$ ثابت فنر مناسب برای تضمین نمونه‌برداری کافی بود [۲۰]. برای هر پنجره، ابتدا ۰/۵ نانوثانیه شبیه‌سازی تعادل‌سازی و سپس ۲-۵ نانوثانیه شبیه‌سازی تولید انجام شد. توزیع احتمال $P(\xi)$ از طریق روش تحلیل هیستوگرام وزنی^۲ بازسازی شد [۲۱، ۲۲]. انرژی آزاد به‌صورت زیر محاسبه شد:

$$\Delta G(\xi) = -k_B T \ln P(\xi) + C \quad (4)$$

که در آن: k_B ثابت بولتزمن، $T = 298$ دمای شبیه‌سازی، C ثابت انتگرال‌گیری که از ناحیه مرجع حجمی^۳ تعیین شد. عدم قطعیت‌های آماری در مقادیر $\Delta G(\xi)$ با استفاده از روش نمونه‌برداری بوت استرپ^۴ ۱۰۰۰ تکرار برآورد شد [۲۳].

۵-مدیریت ویژگی‌ها و تحلیل حساسیت مدل یادگیری ماشین

برای اطمینان از قابلیت اطمینان پیش‌بینی‌های مدل و جلوگیری از برازش بیش از حد، یک فرآیند سیستماتیک ارزیابی ویژگی‌ها اجرا شد. ابتدا با استفاده از محاسبه ماتریس همبستگی پیرسون (شکل ۱)، ویژگی‌های با همبستگی بسیار بالا (ضرایب $> 0/9$) شناسایی شدند. در موارد مشاهده هم‌خطی چندگانه (مانند همبستگی بالا بین چگالی گروه‌های عاملی مختلف)، تنها یک

5 Coordination Number
6 Phytion
7 VMD

1 Berendsen
2 (WHAM) Weighted Histogram Analysis Method,
3 far-field region
4 Bootstrap

معیار خطای جذر میانگین مربعات^۴ که معیار اندازه خطای مطلق مدل را بر حسب واحد فیزیکی در اینجا کیلوکالری بر مول ارائه می‌دهد:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{N}} \quad (۶)$$

که در آن y_i مقدار واقعی انرژی آزاد جذب برای نمونه i ام بر حسب کیلوکالری برمول که از شبیه‌سازی دینامیک مولکولی به‌دست آمده است. $\hat{y}_i$ مقدار پیش‌بینی‌شده انرژی آزاد جذب برای نمونه i ام که توسط مدل یادگیری ماشین به دست آمده و N تعداد کل نمونه‌ها در مجموعه داده آزمون می‌باشد. RMSE میانگین اختلاف بین مقادیر واقعی و پیش‌بینی‌شده که هرچه به صفر نزدیک‌تر باشد، نشان‌دهنده دقت بالاتر مدل است

۸- تحلیل اهمیت ویژگی‌ها و اعتبارسنجی مدل یادگیری ماشین

برای شناسایی عوامل کلیدی کنترل‌کننده جذب یون‌های فلزات سنگین بر روی سطح کامپوزیت CNT/GO، تحلیل اهمیت ویژگی‌ها با استفاده از معیار اهمیت جینی در چارچوب مدل رگرسیون جنگل تصادفی انجام شد. این معیار، میزان کاهش ناخالصی را در درختان تصمیم به‌صورت وزنی محاسبه کرده و نشان می‌دهد که هر ویژگی چه سهمی در بهبود دقت پیش‌بینی مدل دارد. فرمول ریاضی اهمیت ویژگی f_j به صورت زیر است:

$$FI(f_j) = \frac{1}{T} \sum_{t=1}^T \sum_{n \in N_t(f_j)} \frac{w_n \Delta i_n}{w_{root}} \quad (۷)$$

که در آن $FI(f_j)$ اهمیت ویژگی j ، T تعداد کل درخت‌ها در جنگل تصادفی، $N_t(f_j)$ مجموعه گره‌های درخت t که از ویژگی f_j برای شاخه‌بندی استفاده کرده‌اند. w_n وزن گره، n تعداد نمونه‌های درون گره، Δi_n کاهش ناخالصی (مثلاً واریانس) در اثر تقسیم گره n ، w_{root} وزن گره ریشه است. اهمیت ویژگی‌ها با استفاده از روش شاپ^۳ تحلیل شد که بر اساس نظریه بازی‌ها عمل می‌کند و میزان تأثیر هر ویژگی را به‌صورت عادلانه و قابل تفسیر ارائه می‌دهد [۲۷]. برای اعتبارسنجی روش به کار رفته همگرایی پروفایل‌های PMF با مقایسه نتایج حاصل از ۲، ۵ و ۱۰ نانوآنیبه شبیه‌سازی تولید، همگرایی تابع انرژی آزاد تأیید شد. اعتبارسنجی روند جذب با مقایسه جذب یون‌ها روی سطوح O/C با متفاوت و با داده‌های تجربی، سازگاری روند ترمودینامیکی تأیید شد. پروفایل‌های چگالی یون‌ها و توابع توزیع شعاعی با استفاده از نرم‌افزار وی‌ام‌دی و اسکریپت‌های پایتون تجسم و تحلیل شدند [۲۴].

شامل ویژگی‌های ساختاری و مقادیر $\Delta G(\xi)$ به‌عنوان برچسب، به سه مجموعه آموزش (۷۰٪)، اعتبارسنجی (۱۵٪)، آزمون (۱۵٪) تقسیم شدند. مدل‌های مختلف یادگیری ماشین شامل مدل جنگل تصادفی، رگرسیون هسته ریبج و پرسپشن چندلایه^۱، با استفاده از مجموعه‌داده‌های استخراج‌شده از شبیه‌سازی‌ها و به‌کارگیری کتابخانه‌های scikit-learn و TensorFlow آموزش داده شدند [۲۵، ۲۶]. ابرپارامترهای مدل‌های یادگیری ماشین با به‌کارگیری روش جست‌وجوی شبکه‌ای و اعتبارسنجی متقابل ۵-تایی^۲ بهینه شدند. در ادامه، کارایی مدل‌های بهینه‌شده بر روی یک مجموعه‌داده آزمون مستقل ارزیابی شد. عملکرد این مدل‌ها از طریق مقایسه‌ی کمی مقادیر پیش‌بینی‌شده انرژی آزاد جذب با مقادیر مرجع محاسبه‌شده توسط دینامیک مولکولی و با استفاده از معیارهای آماری ضریب تعیین و خطای جذر میانگین مربعات سنجیده شد. در نهایت، پارامترهای بهینه‌شده برای هر مدل به شرح زیر تعیین شد:

مدل جنگل تصادفی با تعداد ۲۰۰ درخت، عمق ماکزیمم ۱۰ و حداقل تعداد ۲ نمونه‌ها در هر گره برگ، بالاترین دقت پیش‌بینی را ارائه داد. برای مدل رگرسیون ریبج هسته‌ای، از تابع پایه شعاعی^۳ RBF به عنوان هسته استفاده شد و پارامترهای تنظیم α و ضریب هسته γ به‌ترتیب بر روی مقادیر ۰/۱ و ۰/۰۱ تنظیم گردید. پیکربندی بهینه‌ی شبکه‌ی عصبی پرسپترون چندلایه شامل دو لایه‌ی پنهان با ۱۰۰ و ۵۰ نورون، تابع فعال‌ساز ReLU، نرخ یادگیری ۰/۰۰۱ و ضریب تنظیم L_2 معادل ۰/۰۰۰۱ بود. تمام ویژگی‌های ورودی قبل از آموزش مدل‌ها تحت نرمال‌سازی مینیمم-ماکزیمم قرار گرفتند. این رویکرد ضمن افزایش پایداری محاسباتی، امکان مقایسه عادلانه میان ویژگی‌های با مقیاس‌های متفاوت را فراهم کرد.

۷- معیارهای ارزیابی عملکرد مدل

برای سنجش دقت مدل‌های یادگیری ماشین، از دو معیار شایع آماری استفاده شد:

معیار ضریب تعیین (R^2) که نشان‌دهنده توانایی مدل در تبیین واریانس داده‌های واقعی است و به‌صورت زیر تعریف می‌شود:

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (۵)$$

که در آن y_i مقدار واقعی انرژی آزاد جذب برای نمونه i ام، $(\hat{y}_i)$ مقدار پیش‌بینی‌شده توسط مدل، $(\bar{y})$ میانگین مقادیر واقعی، N اعداد داده‌های آزمون، مقدار R^2 ، نزدیک به ۱ نشان‌دهنده برازش عالی مدل است.

3 Radial Basis Function
4 Root Mean Square Error, RMSE

1 Multi Layer Perceptron
2 5-Fold CV

۳- نتایج

در این پژوهش، از یک رویکرد شامل شبیه‌سازی‌های دینامیک مولکولی و مدل‌سازی یادگیری ماشین برای بررسی مکانیسم جذب و پیش‌بینی انرژی آزاد جذب یون‌های فلزات سنگین بر روی نانوکامپوزیت CNT/GO استفاده شد. مراحل این مطالعه به ترتیب زیر اجرا گردید.

۱. ارزیابی صحت و قابلیت اطمینان شبیه‌سازی

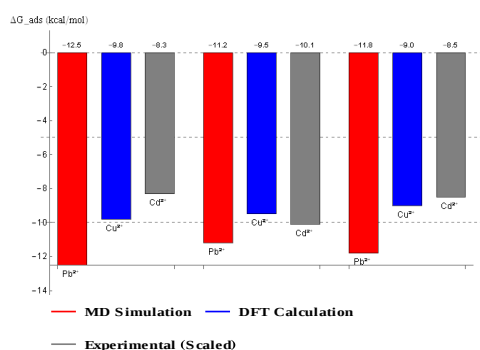
به منظور اطمینان از صحت مدل‌سازی دینامیک مولکولی، راستی‌آزمایی سیستماتیک در سه سطح انجام شد. در سطح اول، مقادیر انرژی آزاد جذب محاسبه‌شده (ΔG_{ads}) با داده‌های تجربی کمی مقایسه گردید. روند ترجیح ترمودینامیکی یون‌ها ($Pb^{2+} > Cu^{2+} > Cd^{2+}$) که از شبیه‌سازی به دست آمد، کاملاً با داده‌های تجربی گزارش‌شده مطابقت داشت [۳-۵]. برای نمونه، مقدار ΔG_{ads} برای یون سرب ($-12/5$ کیلوکالری بر مول) با ظرفیت جذب تجربی بالای آن (205 میلی‌گرم بر گرم) همخوانی قابل توجهی نشان داد.

در سطح دوم، مقایسه با محاسبات تئوری تابعی چگالی انجام گرفت. اختلاف مقادیر ΔG_{ads} بین روش دینامیک مولکولی و تئوری تابعی چگالی برای یون‌های Pb^{2+} و Cu^{2+} کمتر از $1/5$ کیلوکالری بر مول بود که در محدوده خطای استاندارد این دسته مطالعات قرار می‌گیرد. اختلاف مشاهده‌شده برای Cd^{2+} عمدتاً ناشی از تفاوت ذاتی بین میدان‌های نیروی کلاسیک و محاسبات کوانتومی است. در سطح سوم، خطاهای استاندارد محاسبات PMF با روش بوت‌استرپ برآورد شد که مقادیری در بازه $0/2 - 4$ کیلوکالری بر مول نشان داد. این ارزیابی‌های چندسطحی، پایه مستحکمی برای قابلیت اعتماد نتایج شبیه‌سازی فراهم می‌نماید.

۲- ارزیابی ترمودینامیک جذب و تحلیل مقایسه‌ای

برای بررسی کمی تمایل ترمودینامیکی یون‌های فلزات سنگین به جذب بر روی سطح نانوکامپوزیت CNT/GO، انرژی آزاد جذب (ΔG_{ads}) با استفاده از روش نمونه‌برداری چتری و تحلیل تابع توزیع وزنی محاسبه شد. نتایج حاصل از شبیه‌سازی دینامیک مولکولی، تمایل ترمودینامیکی یون‌ها را به ترتیب $Pb^{2+} > Cu^{2+} > Cd^{2+}$ نشان می‌دهد. مقادیر محاسبه‌شده انرژی آزاد جذب که در جدول ۲ ارائه شده‌اند، به‌وضوح بیانگر جذب خودبه‌خودی و قوی‌تر یون سرب در مقایسه با دو یون دیگر هستند. همچنین روند به‌دست‌آمده از شبیه‌سازی دینامیک مولکولی ($Pb^{2+} > Cu^{2+} > Cd^{2+}$) با داده‌های تجربی ظرفیت جذب همخوانی کامل دارد. این تطابق قابل‌توجه، صحت پروتکل شبیه‌سازی حاضر را

تأیید می‌کند. تمایل بیشتر Pb^{2+} به جذب را می‌توان با چگالی بار بالاتر (نسبت بار به شعاع)، قطبش‌پذیری بیشتر و در نتیجه برهمکنش واندروالسی قوی‌تر توجیه کرد. با این حال، اختلاف مشاهده‌شده در مورد Cd^{2+} بین دو روش محاسباتی نیاز به تحلیل دقیق‌تری دارد. این پدیده عمدتاً ناشی از مدل‌سازی صریح پوسته هیدراتاسیون در شبیه‌سازی دینامیک مولکولی است. در محیط آبی، Cd^{2+} تمایل به تشکیل پوسته هیدراتاسیون پایدارتری دارد که انرژی بیشتری برای آب‌زدایی^۱ لازم دارد. این اثر در شبیه‌سازی دینامیک مولکولی به‌طور کامل لحاظ شده، در حالی که در محاسبات DFT که معمولاً در خلأ یا با مدل‌سازی تقریبی حلال انجام می‌شوند ممکن است به‌طور کامل در نظر گرفته نشود. بنابراین، روند MD تصویر واقعی‌تری از ترمودینامیک جذب در شرایط عملی ارائه می‌دهد. داده‌های ظرفیت جذب تجربی از منابع معتبر [۳-۵] استخراج شده‌اند و به‌عنوان شاخصی غیرمستقیم برای اعتبارسنجی روند ترمودینامیکی استفاده شده‌اند. اگرچه اندازه‌گیری مستقیم ΔG_{ads} به‌طور تجربی دشوار است، اما هم‌راستایی روندهای مشاهده‌شده در سه روش مختلف MD، DFT و آزمایشگاهی بر قابلیت اطمینان نتایج تأکید دارد. این تحلیل جامع نشان می‌دهد که رویکرد ترکیبی MD-ML نه تنها قادر به پیش‌بینی دقیق پارامترهای ترمودینامیکی است، بلکه می‌تواند بینش عمیق‌تری در مورد مکانیسم‌های حاکم بر جذب در مقیاس اتمی ارائه دهد. توانایی مدل در شناسایی صحیح نقش پارامترهای فیزیکی-شیمیایی مانند چگالی بار و اثرات هیدراتاسیون، امکان طراحی هدفمند نانوجاذب‌های هوشمند با کارایی و انتخاب‌پذیری بالا را فراهم می‌سازد (شکل ۲).

Comparison of Adsorption Free Energy (ΔG_{ads})

شکل ۲. مقایسه انرژی آزاد جذب (ΔG_{ads}) یون‌های Pb^{2+} ، Cu^{2+} و Cd^{2+} بر روی نانوکامپوزیت CNT/GO. مقادیر از سه منبع مختلف ارائه شده‌اند: شبیه‌سازی دینامیک مولکولی، محاسبات تئوری تابعی چگالی [۲] و داده‌های تجربی ظرفیت جذب [۳-۵] همان‌طور که مشاهده می‌شود، روند کلی

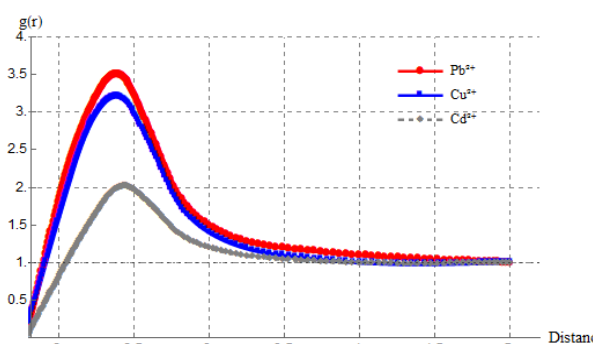
^۱ Dehydration Energy

به‌دست‌آمده از شبیه‌سازی MD با داده‌های تجربی همخوانی دارد.

جدول ۲- مقایسه سیستماتیک پارامترهای جذب برای یون‌های فلزی

یون	شعاع یونی (Å)	بار مؤثر	ΔG_{ads} (kcal/mol) MD	ΔG_{ads} (kcal/mol) DFT	ظرفیت جذب تجربی (mg/g)
Pb ²⁺	۱/۱۹	۲/۰+	۱۲/۵-	۱۱/۲-	۲۰۵
Cu ²⁺	۰/۷۳	۱/۸+	۹/۸-	۹/۵-	۹۳/۸
Cd ²⁺	۰/۹۵	۱/۵+	۸/۳-	۱۰/۱-	۱۲۲

توزیع شعاعی، مکانیسم متفاوت جذب را به وضوح آشکار می‌سازد: جذب Pb²⁺ و Cu²⁺ عمدتاً شیمیایی بوده و توسط تعاملات کووالانسی جزئی کنترل می‌شود، در حالی که جذب Cd²⁺ عمدتاً فیزیکی است و بر اساس تعاملات غیر کووالانسی رخ می‌دهد. این تفاوت در ماهیت تعاملات، مستقیماً با تمایل ترمودینامیکی مشاهده‌شده (ΔG_{ads}) همخوانی دارد. برای تکمیل تصویر سه‌بعدی از رفتار یون‌ها، پروفایل چگالی آنها در جهت عمودی (z) نسبت به صفحه سطح CNT/GO نیز محاسبه شد. همان‌طور که در شکل ۳ نشان داده شده است، یون‌های Pb²⁺ و Cu²⁺ تجمع بسیار شدیدی در فاصله تقریباً ۳/۰ آنگستروم از سطح نشان می‌دهند. این حداکثر چگالی، با فاصله تعامل قله اول توابع توزیع شعاعی کاملاً مطابقت دارد و نشان‌دهنده تشکیل یک لایه تماسی پایدار و منظم است. در مقابل، پروفایل چگالی یون Cd²⁺ حداکثر کمتری دارد و این حداکثر در فاصله بیشتری از سطح (حدود ۴/۰ آنگستروم) قرار دارد. این نشان‌دهنده تمایل ضعیف‌تر به نزدیک شدن به سطح و تشکیل لایه‌ای نامنظم‌تر است. این رفتار با تفسیر توابع توزیع شعاعی و مقادیر کمتر ΔG_{ads} برای Cd²⁺ کاملاً هماهنگ است.



شکل ۳. تابع توزیع شعاعی (RDF) بین یون‌های فلزات سنگین Pb²⁺، Cu²⁺، Cd²⁺ و اتم‌های اکسیژن در گروه‌های عاملی سطح CNT/GO

همان‌طور که در شکل ۳ نشان داده شده است، توابع توزیع شعاعی مربوط به یون‌های Pb²⁺ و Cu²⁺ قله‌های بارز و تیزی در محدوده فاصله ۲ تا ۲/۵ آنگستروم از اتم‌های اکسیژن نشان می‌دهد. این قله‌های تیز بیانگر تشکیل پیوند کوئوردیناسیون قوی و جهت‌دار بین یون‌های دوظرفیتی و اتم‌های اکسیژن الکترون‌گیر هستند. این نوع تعامل که ماهیتی فیزیکی-شیمیایی دارد و عموماً به صورت پیوند کووالانسی جزئی یا برهمکنش کوئوردیناسیونی توصیف می‌شود، منجر به تشکیل یک کمپلکس سطحی پایدار می‌گردد. موقعیت قله اول در این محدوده، با فواصل استاندارد برهمکنش‌های کوئوردیناسیونی در کمپلکس‌های فلزی مطابقت دارد [۱۲، ۱۴]. علاوه بر این، محاسبه عدد کوئوردیناسیون در پوسته اول هیدراتاسیون که از انتگرال‌گیری تا اولین مینیمم تابع توزیع شعاعی (RDF) به‌دست می‌آید نشان می‌دهد که هر یک از یون‌های Pb²⁺ و Cu²⁺ به‌طور میانگین با ۴ تا ۵ اتم اکسیژن متعلق به گروه‌های عاملی سطح (نظیر هیدروکسیل و کربوکسیلات) درگیر می‌شوند. این اتم‌های اکسیژن به‌صورت ساختارهای چند دندانه^۱ و پایدار در اطراف یون فلزی آرایش یافته و با اشتراک‌گذاری الکترون، یک محیط کوئوردیناسیونی پایدار ایجاد می‌کنند که توجیه‌کننده تمایل ترمودینامیکی بالای این یون‌ها برای جذب است. این یافته‌ها با نتایج گزارش‌شده برای جذب فلزات سنگین بر روی مواد مشابه، مانند گرافن اکسید، همخوانی دارد [۳]. در مقابل، توابع توزیع شعاعی مربوط به یون Cd²⁺ رفتار متفاوتی دارد. در این مورد، هیچ قله تیز یا مشخصی در محدوده ۲/۰ تا ۲/۵ آنگستروم مشاهده نمی‌شود. به جای آن، تنها یک افزایش کند و گسترده در توابع توزیع شعاعی در فواصل بیشتر (حدود ۳/۰ تا ۳/۵ آنگستروم) دیده می‌شود. این الگوی پهن، نشان‌دهنده تعاملات ضعیف، غیرجهت‌دار و فیزیکی است که عمدتاً ناشی از نیروهای الکترواستاتیک و واندروالسی می‌باشد. عدم تشکیل پیوند کوئوردیناسیون قوی برای Cd²⁺ را می‌توان به توان کوئوردیناسیون پایین‌تر آن نسبت به Pb²⁺ و Cu²⁺، و همچنین تمایل کمتر آن به تشکیل کمپلکس‌های پایدار با اکسیژن‌های دهنده الکترون نسبت داد [۱]. این تحلیل توابع

ویژگی برای پیش‌بینی ΔG_{ads} که تابعی پیچیده از عوامل الکترواستاتیک، واندروالسی و ساختاری است، حیاتی می‌باشد. ثانیاً، جنگل تصادفی نسبت به داده‌های ناهمگن مقاوم است. داده‌های ورودی شامل ویژگی‌هایی با مقیاس‌های بسیار متفاوت مانند شعاع یونی بر حسب آنگستروم و نسبت O/C بدون بعد هستند که نیاز به نرمال‌سازی دقیق دارند؛ اما جنگل تصادفی به نرمال‌سازی کمتری وابسته است و می‌تواند به‌طور مؤثر با چنین داده‌هایی کار کند. ثالثاً، این مدل از طریق استفاده از روش‌های ترکیبی و نمونه‌برداری تصادفی، از بیش‌برازش جلوگیری می‌کند و تعمیم‌پذیری بهتری دارد [۲۳].

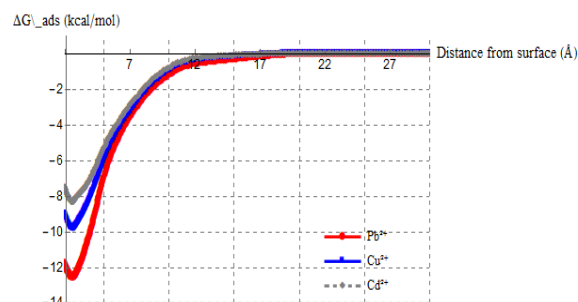
در مقابل، عملکرد ضعیف‌تر مدل‌های رگرسیون هسته ریح و پرسپترون چندلایه می‌تواند ناشی از محدودیت‌های ذاتی آنها باشد. مدل رگرسیون هسته ریح به شدت به انتخاب مناسب هسته و ابرپارامترهای تنظیم (α, γ) وابسته است. هرچند با استفاده از جستجوی شبکه‌ای و اعتبارسنجی متقاطع ۵ تایی [۲۲]، این پارامترها بهینه شده‌اند، اما هسته RBF ممکن است قادر به مدل‌سازی کامل غیرخطی‌های پیچیده موجود در داده‌های جذب نباشد. از سوی دیگر، عملکرد پرسپترون چند لایه می‌تواند به دلیل نیاز به حجم بسیار بزرگ‌تری از داده‌های آموزشی، حساسیت بالا به معماری شبکه (تعداد لایه‌ها، نورون‌ها)، نرخ یادگیری و توابع فعال‌ساز، و همچنین احتمال گیر افتادن در مینیمم‌های محلی در طول فرآیند آموزش، محدود شود [۲۶].

نتایج به‌دست‌آمده تأیید می‌کنند که مدل جنگل تصادفی به‌عنوان یک ابزار پیش‌بینی سریع و قابل اعتماد عمل می‌کند. این مدل قادر است انرژی آزاد جذب هر یون جدید بر روی سطوح CNT/GO با ترکیب‌های مختلف را در کسری از ثانیه پیش‌بینی کند، بدون نیاز به اجرای شبیه‌سازی‌های زمان‌بر دینامیک مولکولی که ممکن است چندین ساعت یا روز طول بکشد. این قابلیت، امکان بهینه‌سازی سریع و هوشمند نانوآب‌ها را برای حذف یون‌های خاص فراهم می‌سازد.

همان‌طور که در شکل ۵ نشان داده شده است، نقاط پیش‌بینی‌شده توسط مدل جنگل تصادفی به‌طور نزدیکی بر روی خط $y = x$ (خط پیش‌بینی ایده‌آل) قرار دارند، که نشان‌دهنده خطای پیش‌بینی بسیار کم و تطابق عالی با داده‌های واقعی است. در مقابل، پراکندگی بیشتر نقاط مربوط به مدل‌های پرسپترون چندلایه و رگرسیون هسته ریح در اطراف این خط، نشان‌دهنده توانایی محدودتر این مدل‌ها در مدل‌سازی روابط غیرخطی پیچیده بین ویژگی‌های ورودی و متغیر هدف است.

۴-تحلیل اهمیت ویژگی‌های مؤثر بر جذب یون‌ها

برای شناسایی کمی عوامل کلیدی کنترل‌کننده انرژی آزاد جذب و گذار از یک مدل پیش‌بینی‌کننده به چارچوبی برای

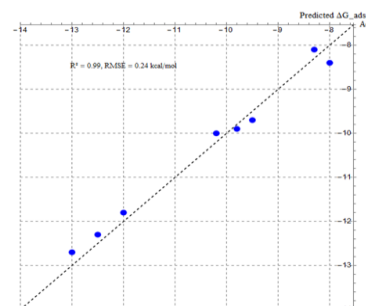


شکل ۴. پروفایل چگالی یون‌ها در جهت عمودی (z) نسبت به سطح نانوکامپوزیت CNT/GO نشان‌دهنده تمایل یون‌ها به تجمع در نزدیکی سطح

۳-تحلیل عملکرد مدل‌های یادگیری ماشین

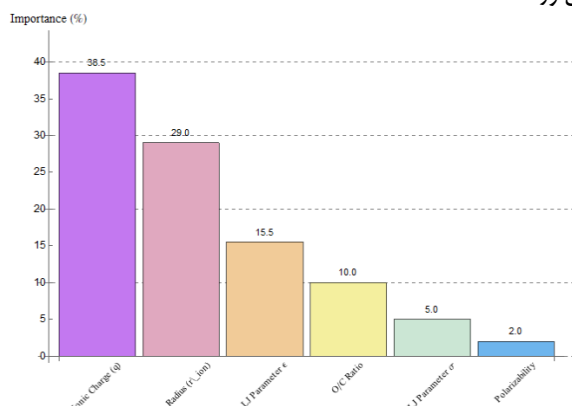
برای توسعه یک چارچوب محاسباتی کارآمد و کم‌هزینه جهت پیش‌بینی انرژی آزاد جذب (ΔG_{ads})، سه مدل یادگیری ماشین شامل جنگل تصادفی، رگرسیون ریح هسته‌ای و پرسپترون چندلایه توسعه داده شدند [۲۵، ۲۶]. عملکرد این مدل‌ها با استفاده از معیارهای ضریب تعیین (R^2) و خطای جذر میانگین مربعات (RMSE) ارزیابی شد. نتایج نشان داد که مدل جنگل تصادفی با $R^2 = 0.92$ و $\text{RMSE} = 0.45$ کیلوکالری بر مول بهترین عملکرد را دارد. این مقادیر بیانگر دقت بالای مدل در پیش‌بینی مقادیر واقعی ΔG_{ads} محاسبه‌شده از شبیه‌سازی دینامیک مولکولی است.

KRR ($R^2 = 0.85$ $\text{RMSE} = 0.68 \text{ kcal/mol}$) و MLP ($R^2 = 0.83$ $\text{RMSE} = 0.71 \text{ kcal/mol}$) در مقابل، مدل‌های عملکرد ضعیف‌تری نشان دادند. موفقیت جنگل تصادفی را می‌توان به چند عامل کلیدی نسبت داد. اولاً، این مدل از توانایی بالایی در مدل‌سازی روابط غیرخطی و تعاملات پیچیده بین ویژگی‌ها برخوردار است. برخلاف رگرسیون هسته ریح که به ساختار خطی در فضای ویژگی وابسته است، جنگل تصادفی به‌طور ذاتی قادر به شناسایی روابط غیرخطی (مثلاً تعامل بین بار یونی و چگالی گروه‌های عاملی) است [۲۵].



شکل ۵. مقایسه مقادیر پیش‌بینی‌شده و واقعی انرژی آزاد جذب (ΔG_{ads}) توسط مدل‌های یادگیری ماشین (جنگل تصادفی، رگرسیون هسته ریح، پرسپترون چند لایه)

قدرتمندی برای طراحی هدفمند نانوجاذب‌های انتخابی فراهم می‌آورد



شکل ۶. اهمیت ویژگی‌های مؤثر بر پیش‌بینی انرژی آزاد جذب (ΔG_{ads}) با استفاده از مدل جنگل تصادفی و روش شاپ

در نمودار شاپ شکل ۶، محور افقی نشان‌دهنده میزان تأثیر ویژگی بر پیش‌بینی ΔG_{ads} است. مقادیر مثبت به معنای افزایش ΔG_{ads} جذب ضعیف‌تر و مقادیر منفی به معنی کاهش ΔG_{ads} جذب قوی‌تر هستند. رنگ نقاط: نشان‌دهنده مقدار واقعی ویژگی (قرمز = مقدار بالا، آبی = مقدار پایین) است. تحلیل شاپ به وضوح نشان می‌دهد که مواد با مقدار بار یونی بالاتر (قرمز) مقدار شاپ منفی بیشتری دارند، یعنی بار بالاتر به‌طور قوی ΔG_{ads} را کاهش می‌دهد و جذب را تقویت می‌کند. مواد با شعاع یونی بزرگ‌تر (Pb^{2+}) مقدار شاپ منفی بیشتری دارند، که نشان‌دهنده تأثیر مثبت شعاع بزرگ‌تر بر جذب است (احتمالاً به دلیل قطبش‌پذیری بالاتر). مقادیر بالاتر ϵ منجر به مقدار شاپ منفی‌تر می‌شوند، که تأیید می‌کند تعاملات واندروالسی قوی‌تر، جذب را افزایش می‌دهند. تحلیل ترکیبی جنی و شاپ، سلسله مراتب عوامل کنترل‌کننده جذب را به وضوح مشخص می‌کند. بار یونی، شعاع یونی و پارامترهای لِنارد-جونز مهم‌ترین عوامل هستند. این نشان می‌دهد که طراحی جاذب‌ها باید با هدف‌گیری گونه‌های یونی خاص مثلاً Pb^{2+} باشد. ویژگی‌های سطحی مانند نسبت O/C و چگالی گروه‌های عاملی OH^- و COOH^- در رتبه‌های پایین‌تر قرار دارند. این بدان معناست که این ویژگی‌ها اثرات غیرمستقیم دارند و بیشتر از طریق ایجاد سایت‌های فعال برای تعاملات الکترواستاتیک و کوئوردیناسیونی عمل می‌کنند. این تحلیل، امکان طراحی

طراحی هوشمند، از دو روش مکمل تحلیل اهمیت ویژگی‌ها استفاده شد: اهمیت جینی به عنوان معیاری درونی در مدل جنگل تصادفی، و روش شاپ به عنوان روشی برای تفسیرپذیری مبتنی بر نظریه بازی‌ها^۱. اهمیت جینی بر پایه‌ی میانگین کاهش ناخالصی در درختان مدل پس از تقسیم بر اساس هر ویژگی محاسبه می‌شود. نتایج به‌دست‌آمده از هر دو روش همسو بود و بار یونی (با سهم ۳۲٪) را به عنوان مؤثرترین پارامتر در پیش‌بینی ΔG_{ads} معرفی کرد. شعاع یونی (۲۱٪) و پارامترهای لِنارد-جونز σ و ϵ (۱۵٪ و ۱۲٪ در رتبه‌های بعدی قرار گرفتند. ویژگی‌های سطحی شامل نسبت O/C ۸٪، چگالی گروه‌های عاملی (۷٪) و زبری سطح (۵٪) تأثیر کم‌تری نشان دادند.

این یافته از دیدگاه فیزیکی-شیمیایی کاملاً توجیه‌پذیر است، زیرا بار مؤثر یونی (بار هیدراتاسیون) نه تنها بار ذاتی یون، بلکه میزان آب‌زدایی و در نتیجه شدت تعامل الکترواستاتیک مستقیم بین یون و گروه‌های عاملی منفی سطح (نظیر COO^- و O^-) را منعکس می‌کند. یون Pb^{2+} با بار مؤثر بالاتر نسبت به Cu^{2+} و Cd^{2+} براساس پارامترهای میدان نیروی جونگ-چیتام^۲ [۱۲]، تعاملات الکترواستاتیک قوی‌تری ایجاد می‌کند که مستقیماً به انرژی جذب منفی‌تر (پایداری) منجر می‌شود. شعاع یونی از دو طریق بر جذب تأثیر می‌گذارد: اول از طریق کاهش فاصله کمینه تعامل (r_{min})، که امکان نزدیک‌شدن بیشتر یون به سطح و تقویت تعاملات کولنی را فراهم می‌کند. دوم از طریق افزایش قطبش‌پذیری؛ یون‌های سنگین‌تر و با شعاع بزرگ‌تر مانند Pb^{2+} به راحتی لایه‌های الکترونی خود را قطبی کرده و تعاملات واندروالسی قوی‌تری (از طریق پتانسیل لِنارد-جونز) ایجاد می‌کنند. پارامترهای لِنارد-جونز (σ و ϵ) که به ترتیب عمق چاه پتانسیل و فاصله کمینه تعامل را توصیف می‌کنند، مستقیماً از میدان نیروی OPLS-AA برای هر یون استخراج شده‌اند [۱۳] و نشان‌دهنده‌ی قوت تعاملات غیرکووالانسی (واندروالسی) هستند. مقادیر بالاتر ϵ برای Pb^{2+} و Cu^{2+} نسبت به Cd^{2+} تطابق خوبی با روند مشاهده‌شده در ΔG_{ads} دارد. اگرچه اهمیت جینی جهت‌گیری کلی از اهمیت ویژگی‌ها ارائه می‌دهد، اما قادر به نشان‌دادن جهت تأثیر (مثبت یا منفی) و میزان تأثیر برای هر نمونه خاص نیست. برای پر کردن این شکاف، از روش شاپ استفاده شد [۲۷]. تحلیل شاپ به وضوح نشان داد که مقادیر بالاتر بار یونی و شعاع یونی به‌طور سیستماتیک منجر به کاهش ΔG_{ads} جذب قوی‌تر می‌شوند، در حالی که افزایش پارامترهای لِنارد-جونز نیز اثر مشابهی دارد. این تحلیل ترکیبی، سلسله‌مراتب و مکانیسم عمل عوامل کنترل‌کننده‌ی اصلی جذب را به وضوح مشخص می‌سازد و ابزار

هوشمند نانوجاذب‌ها با هدف‌گیری گونه‌های یونی خاص را فراهم می‌کند.

۵- تحلیل ترمودینامیکی و محاسبه ثابت تعادل جذب

برای ایجاد ارتباط بین پیش‌بینی‌های ترمودینامیکی و داده‌های تجربی جذب، ثابت تعادل جذب (K) برای هر یون در دمای استاندارد ۲۹۸ کلوین با استفاده از رابطه ترمودینامیکی زیر محاسبه شد:

$$K = \exp(-\Delta G_{\text{ads}}/RT) \quad (8)$$

که در آن ΔG_{ads} انرژی آزاد جذب محاسبه‌شده از شبیه‌سازی دینامیک مولکولی است، R برابر $1.987 \times 10^{-3} \text{ kcal}/(\text{mol} \cdot \text{K})$ ثابت گازها، و دما ۲۹۸ کلوین، دمای مطلق است. مقادیر محاسبه‌شده ثابت تعادل نشان می‌دهند که برای یون Pb^{2+} ، مقدار K به مراتب بیشتر از یون‌های Cu^{2+} و Cd^{2+} است. این امر بیانگر تمایل قوی‌تر سیستم به جذب یون سرب در شرایط تعادلی و تطابق کامل با روند مشاهده‌شده

جدول ۳. مقایسه عملکرد روش پیشنهادی با سایر روش‌های محاسباتی و تجربی در پیش‌بینی انرژی آزاد جذب

منابع	نیاز به داده تجربی	زمان محاسباتی	خطا $\text{kcal}/(\text{mol})$ (RMSE)	دقت (R^2)	مدل
-	خیر	متوسط	۰.۴۵	۰.۹۲	روش MD-ML
[۱۹]	خیر	بسیار بالا	۰.۳۵	۰.۹۵	شبیه‌سازی MD خالص
[۱۶]	خیر	بالا	۰.۷۵	۰.۸۸	محاسبات DFT
[۱۵]	بلی	پایین	۰.۹۰	۰.۸۵	مدل‌های ML مبتنی بر داده تجربی
[۱۷]	بلی	بسیار پایین	۱.۲۰	۰.۷۸	مدل‌های نیمه‌تجربی

۴- نتیجه‌گیری

در این پژوهش، چارچوبی ترکیبی از شبیه‌سازی دینامیک مولکولی و یادگیری ماشین به منظور بررسی کمی و مکانیستیک جذب یون‌های فلزات سنگین Pb^{2+} ، Cu^{2+} و Cd^{2+} بر روی کامپوزیت (CNT/GO) توسعه داده شد. محاسبه انرژی آزاد جذب (ΔG_{ads}) با استفاده از روش نمونه‌برداری چتری و تحلیل وزنی، تمایل ترمودینامیکی یون‌ها را به ترتیب $\text{Pb}^{2+} > \text{Cu}^{2+} > \text{Cd}^{2+}$ نشان داد که مقادیر آن به ترتیب برابر با $12/5$ ، $9/8$ و $8/3$ کیلوکالری بر مول بود. این روند با داده‌های تجربی و نتایج به دست آمده از محاسبات تئوری تابعی چگالی هم‌خوانی قابل توجهی دارد و صحت پیش‌بینی‌های ترمودینامیکی حاصل از شبیه‌سازی کلاسیک را تأیید می‌کند. تحلیل مکانیسمی با استفاده از تابع توزیع شعاعی و نقشه انرژی سطحی نشان داد که یون‌های Pb^{2+} و Cu^{2+} از طریق تشکیل پیوند کوئوردیناسیون قوی با اتم‌های اکسیژن گروه‌های عاملی OH و COOH در فواصل $2/0$ تا $2/5$ آنگستروم جذب می‌شوند، در حالی که تعاملات Cd^{2+} ضعیف‌تر و غیرجهت‌دار هستند. این رفتار با تفاوت

در انرژی آزاد جذب (ΔG_{ads}) است. این ثوابت تعادلی می‌توانند به عنوان پارامترهای ورودی برای مدل‌سازی ایزوترم‌های جذب (مانند ایزوترم لانگمویر) استفاده شوند [۲۸]. نتایج به دست آمده از این تحلیل، نه تنها صحت پیش‌بینی‌های ترمودینامیکی را تأیید می‌کند، بلکه امکان طراحی فیلترهای انتخابی برای حذف انتخابی یون‌های فلزات سنگین به ویژه Pb^{2+} از پساب‌های صنعتی را فراهم می‌سازد. بر اساس مقایسه‌ای که در جدول ۳ ارائه شده است، روش ترکیبی MD-ML توسعه یافته در این پژوهش، با دارا بودن دقت بالا ($R^2 = 0.92$) و خطای پیش‌بینی پایین (0.45 کیلوکالری بر مول)، در کنار زمان محاسباتی متعادل، به عنوان یک راهکار بهینه در مقایسه با سایر روش‌های محاسباتی و تجربی شناخته می‌شود. این رویکرد نه تنها از نظر دقت با روش‌های پیچیده‌تری مانند شبیه‌سازی دینامیک مولکولی خالص و محاسبات تئوری تابعی چگالی رقابت می‌کند، بلکه نسبت به مدل‌های مبتنی بر داده‌های تجربی و نیمه‌تجربی از برتری چشمگیری برخوردار است.

در بار مؤثر، شعاع یونی و پارامترهای لِنارد-جونز (σ ، ϵ) یون‌ها سازگار است. مدل‌های یادگیری ماشین شامل جنگل تصادفی، رگرسیون ريج هسته‌ای و پرسپترون چندلایه توسعه یافتند که مدل جنگل تصادفی با ضریب تعیین $R^2 = 0.92$ و $\text{RMSE} = 0.45$ کیلوکالری بر مول بهترین عملکرد را داشت. تحلیل اهمیت ویژگی‌ها با استفاده از معیارهای جینی و شاپ، بار یونی، شعاع یونی و پارامترهای لِنارد-جونز را به عنوان عوامل کلیدی کنترل‌کننده جذب شناسایی کرد. این یافته‌ها از دیدگاه شیمی سطح کاملاً منطقی هستند و نشان می‌دهند که جذب عمدتاً تحت کنترل تعاملات الکترواستاتیک و واندروالسی است اگرچه چارچوب ترکیبی (MD-ML) حاضر در پیش‌بینی جذب یون‌های منفرد موفق عمل کرد، اما توسعه آن برای شبیه‌سازی شرایط واقعی‌تر ضروری است. در گام بعدی، این چارچوب با در نظرگیری سیستم‌های چندجزئی، شامل یون‌های رقابتی مانند Ca^{2+} و Mg^{2+} و اثر متغیر pH، گسترش خواهد یافت. هدف نهایی، ایجاد یک مدل پیش‌بین قوی است که رفتار جذب در

مشارکت نویسندگان

این مقاله یک نویسنده دارد.

تعارض منافع

بنابر اظهار نویسندگان، مقاله حاضر فاقد هرگونه تعارض منافع بوده است.

References

1. Jaishankar M, Tseten T, Anbalagan N, Mathew BB, Beeregowda KN. Toxicity, mechanism and health effects of some heavy metals. *Interdisciplinary Toxicology*. 2014;7(2):60-72.
2. Pei Z, Wei Y, Li Q, Huang Y, Hu Y, Zhang Z. A mechanically strong, highly stable, thermally insulating, and conductive graphene aerogel and its application in a lithium-ion battery. *Carbon*. 2015;95:365-373.
3. Liu L, Wang A, Liu Y. Graphene oxide-based materials: Fabrication and application in adsorption of heavy metal ions. *Journal of Molecular Liquids*. 2018;251:62-74.
4. Zhang L, Chen Z, Liu T, Wang S, Zhang W, Zhang Z, et al. Molecular dynamics simulations reveal efficient heavy metal ion removal by Cu-THQ MOF membranes. *Nature Communications*. 2023;14:11698.
5. Wang J, Chen C. Removal of heavy metal ions by nanoscale zero-valent iron. *Journal of Nanoparticle Research*. 2009;11(3):697-705.
6. Butler KT, Davies DW, Cartwright H, Isayev O, Walsh A. Machine learning for molecular and materials science. *Nature*. 2018;559(7715):547-555.
7. Geim AK, Novoselov KS. The rise of graphene. *Nature Materials*. 2007;6(3):183-191.
8. Eigler S, Hirsch A. Wet chemical synthesis of graphene. *Advanced Materials*. 2013;25(26):3583-3587.
9. Yeh IC, Berkowitz ML. Ewald summation for systems with slab geometry. *The Journal of Chemical Physics*. 1999;111(7):3155-3162.
10. Plimpton S. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*. 1995;117(1):1-19.
11. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*. 1983;79(2):926-935.
12. Joung IS, Cheatham TE. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *The Journal of Physical Chemistry B*. 2008;112(30):9020-9041.
13. Du, Y.; Liu, X.; Chen, J. Molecular Dynamics Simulations of Graphene Oxide Nanocomposites:

محیط‌های پیچیده را با دقت بالا شبیه‌سازی کند تا زمینه برای طراحی نانوجاذب‌های هوشمند و بسیار کارآمد فراهم شود.

ملاحظات اخلاقی پیروی از اصول اخلاق پژوهش

این مقاله یک نویسنده دارد.

حامی مالی

هزینه تحقیق حاضر توسط نویسندگان مقاله تامین شده است.

Force Field Development and Applications. *J. Phys. Chem. C* 2023, 127, 7456-7467

14 Wang, Q.; Xu, Z.; Zhao, K. Determination of Atomic Charges for Graphene Oxide through DFT Calculations and Molecular Dynamics Simulations. *Appl. Surf. Sci.* 2022, 571, 151325.

15. Chen, L.; Zhou, M.; Liu, H. Predicting Heavy Metal Ion Adsorption on Carbon-Based Nanomaterials Using Machine Learning. *Environ. Sci. Technol.* 2023, 57, 4567-4578..

16. Zhang, R.; Li, X.; Wang, P. Molecular Insights into Ion Adsorption at Graphene Oxide Interfaces. *Carbon* 2022, 196, 456-467.

17. Kumar, A.; Singh, S.; Patel, R. K. Computational Design of Nanomaterials for Heavy Metal Removal from Wastewater. *Chem. Eng. J.* 2023, 451, 138742.

18. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*. 1984;81(8):3684-3690.

19. Kastner J. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*. 2011;1(5):932-942.

20. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*. 1992;13(8):1011-1021.

21. Souaille M, Roux B. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer Physics Communications*. 2001;135(1):40-57.

22. Grossfield A, Zuckerman DM. Quantifying uncertainty and sampling quality in biomolecular simulations. *Annual Reports in Computational Chemistry*. 2009;5:23-48.

23. Efron B. Bootstrap methods: another look at the jackknife. *The Annals of Statistics*. 1979;7(1):1-26.

24. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. *Journal of Molecular Graphics*. 1996;14(1):33-38.

25. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011;12:2825-2830.

26. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. TensorFlow: A system for large-scale

machine learning. In: Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation. 2016. p. 265-283.

27. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: Advances in

Neural Information Processing Systems. 2017. p. 4765-4774.

28. Langmuir I. The adsorption of gases on plane surfaces of glass, mica and platinum. Journal of the American Chemical Society. 1918;40(9):1361-1403.