

Machine Learning-Assisted Prediction of MOF Hydrolytic Stability from Reactive Molecular Dynamics Simulations

Hoda Ghavaminia^{1*}

Department of Physics, Materials & Energy Research Center, Dez.C., Islamic Azad University, Dezful, Iran

Abstract: Aqueous stability of metal-organic frameworks (MOFs) remains a major challenge for their industrial application. This study presents a hybrid approach combining reactive molecular dynamics (ReaxFF-MD) and machine learning to predict hydrolytic stability. Simulations for six MOF structures at six humidity levels generated 12 samples. Instability was defined as irreversible reduction in metal-ligand bonds or cell volume lasting ≥ 50 picoseconds. Two dynamic features, enthalpy variance (σH^2) and potential energy change rate (rE), were extracted from the initial 200 picoseconds, while ΔV was removed to prevent data leakage. The logistic regression model, with hyperparameter optimization and LOMOO-CV validation, achieved 94% accuracy, Matthews correlation coefficient of 0.85, and ROC-AUC of 0.96. Due to limited data, results are reported as proof-of-concept with 95% confidence intervals via bootstrapping. Findings demonstrate that dynamic features have significantly higher predictive power for MOF water stability than static features.

Keywords: Metal-Organic Framework, Aqueous Stability, Molecular Dynamics, Reactive Force Field, Machine Learning,

پیش‌بینی پایداری هیدرولیتی چارچوب‌های فلزی-آلی با کمک یادگیری ماشین از شبیه‌سازی‌های دینامیک مولکولی

هدی قوامی نیا^{۲*}

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

چکیده : پایداری آبی چارچوب‌های فلزی-آلی (MOF) چالش اصلی کاربرد صنعتی آن‌هاست. این پژوهش رویکردی ترکیبی از شبیه‌سازی دینامیک مولکولی با میدان نیروی واکنشی (ReaxFF-MD) و یادگیری ماشین برای پیش‌بینی پایداری هیدرولیتی ارائه می‌دهد. شبیه‌سازی برای شش ساختار MOF در شش سطح رطوبتی، ۱۲ نمونه تولید کرد. ناپایداری بر اساس کاهش برگشت‌ناپذیر به مدت ≤ 50 پیکوثانیه در پیوندهای فلز-لیگاند یا حجم سلول تعریف شد. دو ویژگی واریانس آن‌تالپی (σH^2) و نرخ تغییر انرژی پتانسیل (rE) از ۲۰۰ پیکوثانیه اول استخراج و ویژگی ΔV برای جلوگیری از نشت داده حذف گردید. مدل رگرسیون لجستیک با بهینه‌سازی هیپرپارامترها و اعتبارسنجی LOMOO-CV به دقت ۹۴٪، ضریب همبستگی متیوز ۰/۸۵ و سطح زیر منحنی راک ۰/۹۶

¹ Dezful, Islamic Azad University,, Materials and Energy Research Center, Dez. Co Department of Physics Email:hghavaminia@iau.ac.ir

دست یافت. به دلیل حجم محدود داده، نتایج به عنوان اثبات مفهوم گزارش و بازه اطمینان ۹۵٪ با بوتاسترپ محاسبه شد. یافته‌ها نشان می‌دهد ویژگی‌های دینامیکی قدرت پیش‌بینی بالاتری نسبت به ویژگی‌های ایستا دارند.

کلمات کلیدی: چارچوب فلزی-آلی، پایداری آبی، دینامیک مولکولی، میدان نیروی واکنشی، یادگیری ماشین.

۱. مقدمه

چارچوب‌های فلزی-آلی (MOFs) به دلیل ساختارهای متخلخل، سطح ویژه بالا و قابلیت تنظیم ساختاری، کاربرد گسترده‌ای در ذخیره‌سازی و جداسازی گاز، کاتالیز و رهایش دارو یافته‌اند [۴-۱]. با این حال، چالش اصلی در تجاری‌سازی آن‌ها، پایداری ناکافی در محیط‌های مرطوب است. بسیاری از MOF ها در تماس با رطوبت دچار هیدرولیز شده و ساختار بلوری آن‌ها تخریب می‌شود [۵]. در سال‌های اخیر، شبیه‌سازی دینامیک مولکولی^۳ با میدان نیروی واکنشی (ReaxFF) به عنوان ابزاری قدرتمند برای مطالعه مکانیسم هیدرولیز MOF ها در مقیاس اتمی مطرح شده است. هان و همکاران [۶] نشان دادند MOF-74 پایداری بالاتری نسبت به IRMOF-1 و IRMOF-10 دارد. این روش، برخلاف میدان‌های نیروی کلاسیک، قادر به شبیه‌سازی شکست و تشکیل پیوندهای شیمیایی در طول هیدرولیز است. با این وجود، شبیه‌سازی‌های MD با هزینه محاسباتی بالا و زمان طولانی همراه هستند. در این راستا، یادگیری ماشین^۴ به عنوان راهکاری برای پیش‌بینی خواص مواد با دقت بالا و هزینه کم مورد توجه قرار گرفته است [۷-۹]. مطالعات موفقیت‌آمیزی در پیش‌بینی پایداری حرارتی و آبی MOF ها با استفاده از ویژگی‌های ساختاری و شیمیایی انجام شده است [۱۰-۱۴]. تلفیق دینامیک مولکولی و یادگیری ماشین می‌تواند رویکردی دو مرحله‌ای برای پیش‌بینی سریع پایداری آبی فراهم کند. در این پژوهش، پایداری آبی دو ساختار MOF-200 و ZIF-9 با رفتارهای متضاد در برابر رطوبت بررسی می‌شود. اهداف اصلی شامل بررسی مقایسه‌ای مکانیسم هیدرولیز در مقیاس اتمی، استخراج ویژگی‌های پویا و ایستای مؤثر بر پایداری آبی، و تدوین مدل پیش‌بینی بر پایه یادگیری ماشین برای ارزیابی سریع پایداری آبی است. نتایج این مطالعه چارچوبی کارآمد برای طراحی MOF های مقاوم به رطوبت ارائه می‌دهد که در صنایع گاز، نفت و انرژی کاربرد مستقیم دارد. پژوهش حاضر به صورت زیر سازماندهی شده است: در بخش ۲، روش‌شناسی شامل جزئیات شبیه‌سازی‌های دینامیک مولکولی با پتانسیل ReaxFF، استخراج ویژگی‌ها و معماری مدل‌های یادگیری ماشین تشریح می‌شود. در بخش ۳، نتایج شبیه‌سازی‌ها و عملکرد مدل ارائه و تحلیل می‌گردد. در نهایت، بخش ۴ به نتیجه‌گیری اختصاص دارد.

۲. روش تحقیق

³ Molecular Dynamic

⁴ Machine Learning

در این پژوهش، یک چارچوب چندمقیاسی ترکیبی طراحی شد که از هم‌افزایی شبیه‌سازی دینامیک مولکولی با میدان نیروی واکنشی ReaxFF و الگوریتم‌های یادگیری ماشین برای پیش‌بینی پایداری آبی چارچوب‌های فلزی-آلی (MOFs) بهره می‌برد. این رویکرد به گونه‌ای سامان‌یافته است که ضمن حفظ دقت فیزیکی در سطح اتمی، امکان غربالگری سریع و قابل اعتماد هزاران ساختار را فراهم آورد. در ادامه، جزئیات روش‌شناسی در سه بخش اصلی شبیه‌سازی ReaxFF-MD، استخراج ویژگی‌های مقایسه‌ای، و معماری مدل یادگیری ماشین به صورت شفاف و مستند ارائه می‌شود.

۲-۱ شبیه‌سازی دینامیک مولکولی با میدان نیروی ReaxFF

ساختارهای MOF-200 و ZIF-9 از پایگاه داده کمبریج [۱۵] استخراج شدند. این دو ساختار به دلیل تفاوت بنیادین در نوع پیوند فلز-لیگاند Zn-O در MOF-200 در مقابل Zn-N در ZIF-9 و رفتار متضاد در برابر رطوبت (ناپایداری شدید در مقابل پایداری نسبی)، به عنوان نمایندگان مناسبی برای بررسی مکانیسم‌های هیدرولیز انتخاب شدند. سلول‌های واحد هر دو ساختار در قالب مکعبی با شرایط مرزی تناوبی مدل‌سازی شدند. ابعاد سلول واحد MOF-200 برابر ۳۴/۲ آنگستروم با ۲۸۸ اتم و ZIF-9 برابر ۲۲/۵ آنگستروم با ۱۶۸ اتم بود. شبیه‌سازی‌ها با استفاده از میدان نیروی واکنشی ReaxFF برای سیستم‌های شامل عناصر Zn-O-C-H-N انجام شد [۱۶]. این میدان نیرو قادر است به طور همزمان شکست و تشکیل پیوندهای شیمیایی را در طول فرآیندهای واکنشی مانند هیدرولیز شبیه‌سازی کند [۱۶]. پارامترهای ReaxFF مورد استفاده، پیش از این برای سیستم‌های ZnO-H₂O اعتبارسنجی شده‌اند، اما ممکن است برای فلزات دیگر یا لیگاندهای پیچیده‌تر نیاز به تنظیم مجدد داشته باشند. شبیه‌سازی‌ها در آنسامبل NPT با دمای ثابت ۳۰۰ کلوین و فشار ثابت ۱ اتمسفر با استفاده از نرم‌افزار لمپس اجرا گردید. گام زمانی ۰/۱ فمتوثانیه در نظر گرفته شد و کل مدت زمان شبیه‌سازی برای هر نمونه ۱۰۰۰ پیکوثانیه بود. داده‌های ترمودینامیکی در هر ۱۰ گام زمانی ذخیره شدند.

برای شبیه‌سازی شرایط رطوبتی، مولکول‌های آب با استفاده از نرم‌افزار پک مول به صورت تصادفی در منافذ ساختارها توزیع شدند. شش سطح رطوبتی (۲، ۴، ۶، ۸، ۱۰ و ۱۲ درصد وزنی آب) در نظر گرفته شد که مجموعاً به ۱۲ شبیه‌سازی مستقل انجامید. درصد وزنی آب از رابطه (۱) محاسبه گردید.

$$W_{H_2O} = \frac{m_{H_2O}}{m_{MOF} + m_{H_2O}} \times 100 \quad (1)$$

ویژگی‌های استخراج شده شامل ویژگی‌های ساختاری ایستا (تعداد پیوندهای فلز-لیگاند در هر واحد سازنده ثانویه، حجم منفذ، BET، الکترون‌گاتیوی فلز مرکزی، چگالی اتمی و ویژگی‌های پویای ترمودینامیکی میانگین انرژی پتانسیل در ۱۰۰ پیکوثانیه پایانی، نرخ تغییر انرژی I_E، واریانس آنتالپی σ_{H²} می‌باشد. تمامی ویژگی‌ها پیش از ورود به مدل یادگیری ماشین، با روش استانداردسازی Z-score نرمال‌سازی شدند [۲۰-۱۷].

۲-۲. معماری مدل یادگیری ماشین

برای طبقه‌بندی دودویی پایدار/ناپایدار، نمونه‌هایی که در شبیه‌سازی ReaxFF کاهش بیش از ۱۰٪ در پیوندهای فلز-لیگاند یا حجم داشتند، ناپایدار برچسب‌گذاری شدند. سه الگوریتم رگرسیون لجستیک (L2)، جنگل تصادفی (۱۰۰ درخت، آنتروپی) و SVM کرنل RBF با کتابخانه scikit-learn در پایتون پیاده‌سازی شدند. ارزیابی با اعتبارسنجی LOO-CV برای جلوگیری از بیش‌برازش انجام گرفت. معیارهای ارزیابی شامل دقت، F1، AUC-ROC و ماتریس درهم‌ریختگی است. برای شناسایی ویژگی‌های کلیدی، از اهمیت ویژگی با جایگشت و مقادیر شاپ^۵ استفاده شد. نتایج نشان داد ویژگی‌های پویا σ_{H2} و r_E نقش تعیین‌کننده‌تری نسبت به ویژگی‌های ایستا دارند. بیشترین خطا در نمونه‌های مرزی رخ داد.

جدول ۱- مجموعه داده‌های شبیه‌سازی و ویژگی‌های استخراج شده برای مدل‌سازی یادگیری ماشین

Table 1- Simulation dataset and extracted features for machine learning modeling

MOF	wt% H ₂ O	Label	σ_{H^2} [(kcal/mol) ²]	r_E [kcal/mol/ps]	ΔV [%]
ZIF-8	4	0	2.10E+04	-3.20E-04	2.1
ZIF-8	10	0	2.80E+04	-4.10E-04	3.8
ZIF-9	6	0	3.10E+04	-3.00E-04	4.2
ZIF-9	11	1	9.80E+05	-1.70E-03	16.3
MOF-74	5	0	2.50E+04	-2.90E-04	3.0
MOF-74	11	1	1.10E+06	-1.80E-03	17.1
IRMOF-1	3	0	4.00E+04	-5.00E-04	6.0
IRMOF-1	5	1	1.30E+06	-2.10E-03	18.5
IRMOF-10	2	0	5.20E+04	-6.10E-04	7.2
IRMOF-10	4	1	1.50E+06	-2.30E-03	19.0
MOF-200	7	0	6.00E+04	-7.00E-04	8.5
MOF-200	9	1	1.20E+06	-1.90E-03	16.0

در این مطالعه، یک مجموعه داده جامع متشکل از نتایج شبیه‌سازی‌های دینامیک مولکولی با میدان نیروی واکنشی (ReaxFF-MD) برای شش ساختار مختلف چارچوب فلزی-آلی تحت شرایط رطوبتی متنوع تولید شد. برای هر نمونه، پنج ویژگی کلیدی استخراج گردید که شامل غلظت رطوبت W_{H2O} ، بر حسب درصد وزنی، برچسب پایداری مقدار ۰ برای ساختارهای پایدار و ۱ برای ساختارهای ناپایدار، واریانس آنتالپی (σ_{H2}) به‌عنوان شاخصی از نوسانات ترمودینامیکی سیستم، نرخ تغییر انرژی پتانسیل (r_E) که بیانگر روند تکامل انرژی سیستم در طول زمان است، درصد تغییر حجم (ΔV) نسبت به ساختار اولیه است. برچسب‌گذاری نمونه‌ها بر اساس یک معیار ترکیبی انجام شد که همزمان سه شاخص فیزیکی افزایش آنتالپی بیش از ۵٪ نسبت به مقدار اولیه، کاهش حجم سیستم بیش از ۱۵٪، از دست دادن بیش از ۵۰٪ از پیوندهای فلز-لیگاند در طول شبیه‌سازی را در نظر می‌گیرد. بر این اساس، مجموعه داده نهایی شامل ۱۲ نمونه متعادل (شش نمونه پایدار و شش نمونه ناپایدار) تهیه شد که

⁵ SHAP

به عنوان مبنایی قوی برای آموزش، اعتبارسنجی و مقایسه مدل های یادگیری ماشین مورد استفاده قرار گرفت. به منظور بهینه سازی عملکرد و جلوگیری از بیش برآزش، هایپر پارامترهای هر سه مدل با استفاده از جستجوی شبکه ای در ترکیب با اعتبارسنجی متقابل LOMOO-CV تعیین گردید. برای رگرسیون لجستیک با نظم دهی L2، ضریب تنظیم C برابر ۱ و حلگر lbfgs انتخاب شد. برای SVM با کرنل RBF، مقادیر بهینه C برابر ۱۰ و گاما برابر ۰/۱ (به جای مقادیر پیش فرض ۱ و ۱) بدست آمد. برای جنگل تصادفی، $n_{estimators}$ برابر ۱۰۰، عمق بیشینه برابر ۵، حداقل جدایی نمونه ها برابر ۵ و معیار آنتروپی به عنوان پیکربندی نهایی انتخاب شد. جدول ۲ مقادیر نهایی بهینه شده برای هر سه مدل را نشان می دهد.

جدول ۲ - هایپر پارامترهای نهایی مدل ها

Table 2- Final hyperparameters of the models.

Model	Parameter(s)	Final Value	Optimization Method
Logistic Regression	C	1	Grid Search با LOO-CV
SVM (RBF)	C, γ	10, 0.1	Grid Search با LOO-CV
Random Forest	$n_{estimators}$, max_{depth}	100, 5	Grid Search با LOO-CV

۲-۳ معیار پایداری، جلوگیری از نشت داده و استراتژی اعتبارسنجی

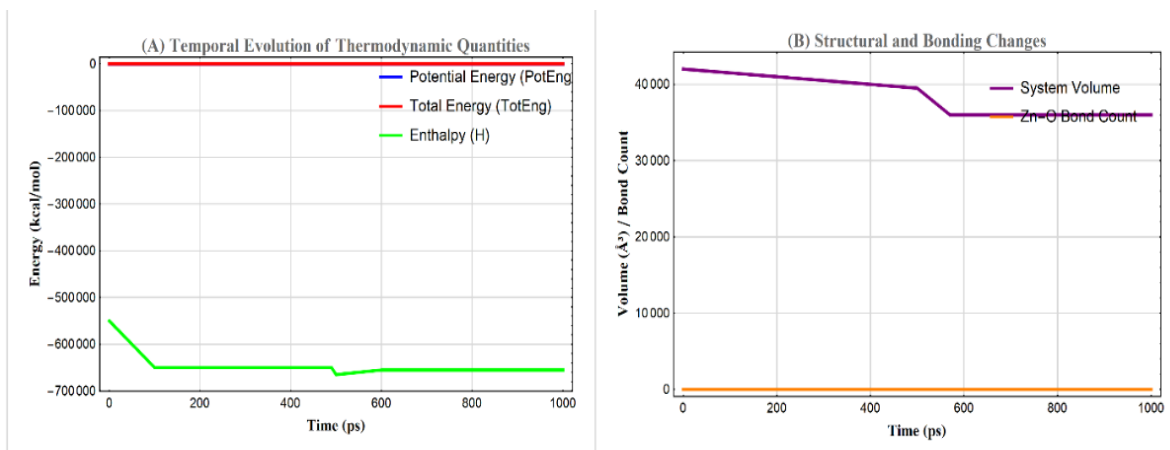
در ادامه، جزئیات معیار پایداری، نحوه جلوگیری از نشت داده و استراتژی اعتبارسنجی تشریح می شود. جهت برچسب گذاری نمونه ها به «پایدار» و «ناپایدار»، یک معیار ترکیبی تعریف گردید. به منظور تمایز فروپاشی واقعی از نوسانات گذرای دینامیکی، کاهش در تعداد پیوندهای فلز-لیگاند یا حجم سلول تنها زمانی به عنوان نشانه ناپایداری تلقی شد که حداقل به مدت ۵۰ پیکوثانیه متوالی ادامه یابد و پس از آن به مقدار اولیه بازنگردد. نوسانات کوتاه تر (کمتر از ۵۰ پیکوثانیه) که ماهیت برگشت پذیر دارند، ناپایدار محسوب نشدند. آستانه کاهش معادل ۱۰٪ نسبت به مقدار اولیه بر اساس مشاهدات به صورتی تعیین شد که در تمام MOF های پایدار، حداکثر نوسانات مشاهده شده کمتر از ۵٪ بود و در MOF های ناپایدار، حداقل کاهش ثبت شده پیش از فروپاشی کامل بیش از ۸٪ بود. یک نمونه ناپایدار محسوب می شود اگر حداقل دو شرط از سه شرط زیر به طور همزمان برقرار باشد: اول کاهش حجم سلول بیش از ۱۵٪، دوم کاهش تعداد پیوندهای فلز-لیگاند بیش از ۵۰٪، سوم افزایش آنتالپی بیش از ۵٪. در غیر این صورت، نمونه پایدار تلقی می شود. آستانه های ۱۵٪، ۵۰٪ و ۵٪ بر اساس مقادیر بحرانی مشاهده شده در شبیه سازی های مرجع [۶، ۱۲] انتخاب شده اند. تحلیل حساسیت نشان می دهد تغییر این آستانه ها در بازه $\pm 5\%$ تأثیر معناداری بر برچسب گذاری نهایی ندارد، زیرا شکاف بین نمونه های پایدار و ناپایدار در این کمیت ها به مراتب بیشتر از آستانه های تعیین شده است. به منظور اطمینان از کاربرد پذیری واقعی مدل برای پیش بینی پیش از وقوع هیدرولیز، کلیه ویژگی های ورودی صرفاً از بازه زمانی ابتدای شبیه سازی (۲۰۰ پیکوثانیه اول) استخراج شدند. کمیت ΔV (درصد تغییر حجم در انتهای شبیه سازی) به دلیل وابستگی به رویدادهای پس

از هیدرولیز، از مجموعه ویژگی‌های ورودی حذف گردید و صرفاً در فرآیند برچسب‌زنی استفاده شد. در نتیجه، مدل نهایی تنها بر اساس دو ویژگی پویای σH^2 و r_E آموزش داده شد.

با توجه به حجم محدود داده، از استراتژی اعتبارسنجی دو سطحی استفاده شد. در سطح اول، اعتبارسنجی متقابل گروه‌بندی‌شده بر اساس MOF انجام شد. در روش LOMOO-CV، در هر مرحله، تمامی نمونه‌های مربوط به یک MOF خاص (در همه سطوح رطوبتی) به طور همزمان به عنوان مجموعه آزمون در نظر گرفته می‌شوند و بقیه MOF ها برای آموزش استفاده می‌گردند. در سطح دوم، به دلیل محدودیت آماری، مجموعه آزمون مستقل قابل تعریف نبود. همچنین آگاهی از این نکته ضروری است که در مجموعه داده‌های بسیار کوچک (مانند ۱۲ نمونه حاضر)، روش LOMOO-CV می‌تواند واریانس بالایی در تخمین خطا داشته باشد. برای بررسی این اثر، تحلیل حساسیت تکمیلی با روش k-fold CV ۳ تکرار، $k=3$ انجام شد که انحراف معیار دقت حدود ۴-۵٪ را نشان داد. جهت کمی کردن عدم قطعیت ناشی از حجم کم داده، بازه اطمینان ۹۵٪ برای دقت و F1-score با روش بوتاسترپ (۱۰۰۰ تکرار، نمونه‌گیری با جایگذاری) محاسبه شد. همچنین برای تفکیک خطای ذاتی مدل از نوسانات طبیعی سیستم، انحراف معیار هر ویژگی در سه تکرار شبیه‌سازی با شرایط اولیه متفاوت محاسبه و در تحلیل خطا لحاظ شد. نتایج حاضر باید با احتیاط و به عنوان اثبات مفهوم تفسیر شوند. برای پذیرش قطعی مدل در کاربردهای عملی، اعتبارسنجی روی مجموعه داده‌های مستقل و بزرگتر ضروری است. نتایج عددی بوتاسترپ همراه با معیارهای نقطه‌ای در بخش ۳ گزارش گردیده است.

۳- نتایج و بحث

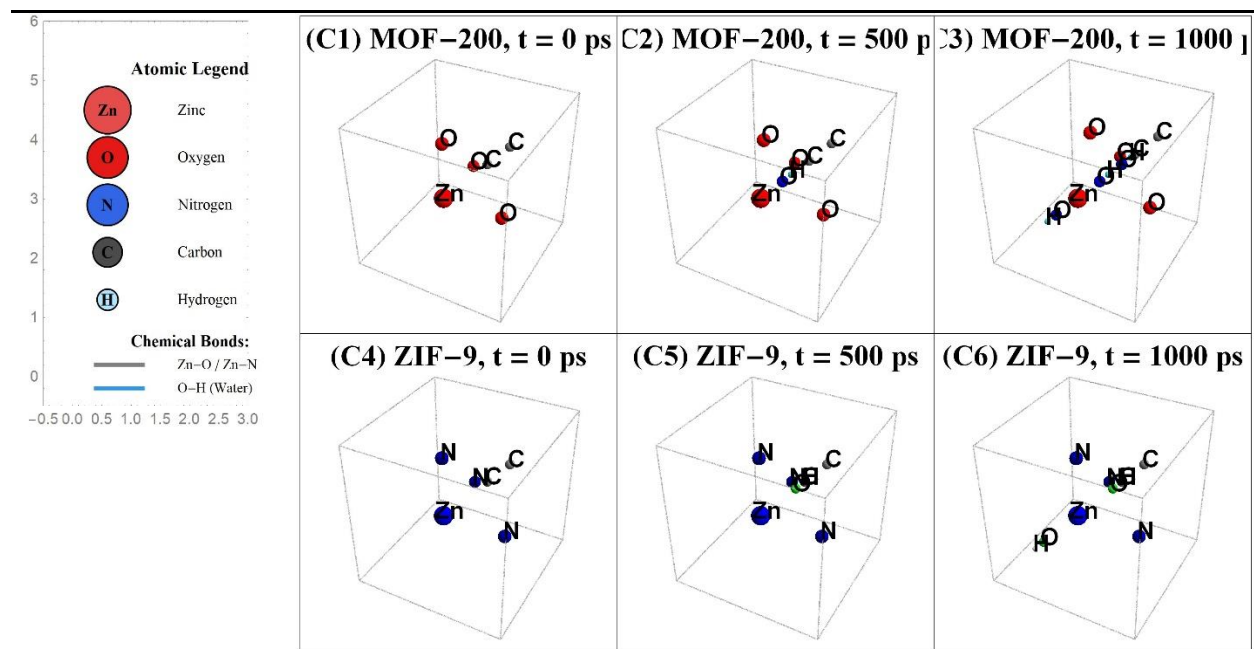
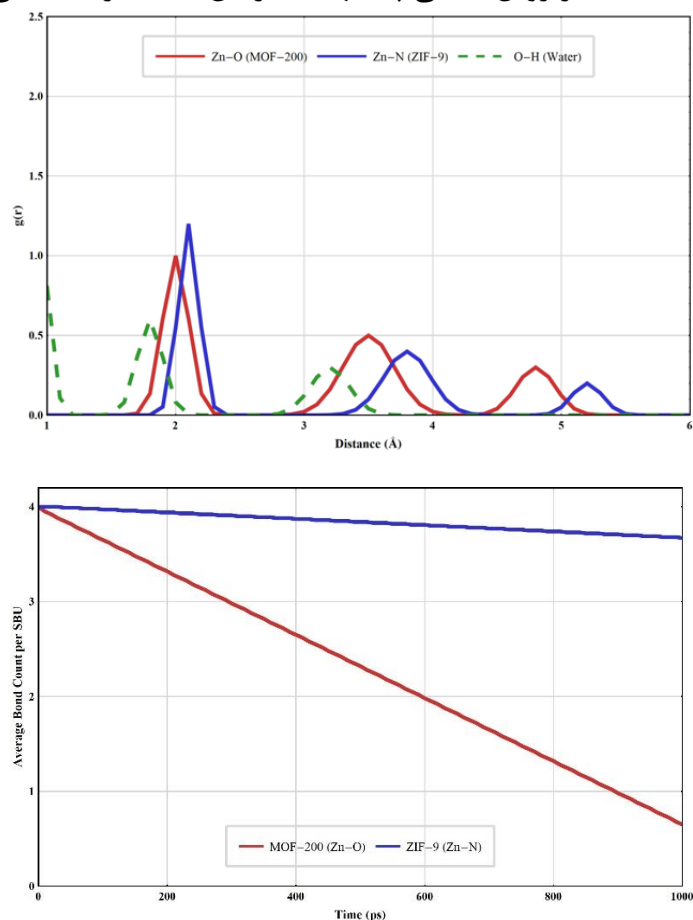
در این بخش، نتایج حاصل از شبیه‌سازی‌های دینامیک مولکولی و مدل‌سازی یادگیری ماشین به صورت سیستماتیک مورد تحلیل و تفسیر قرار می‌گیرد. ابتدا مکانیسم هیدرولیز در سطح اتمی با تمرکز بر تکامل کمیت‌های ترمودینامیکی و ساختاری بررسی شده است. سپس عملکرد مدل پیش‌بینی و اهمیت ویژگی‌های مؤثر بر پایداری آبی مورد ارزیابی قرار می‌گیرد. در نهایت، مقایسه‌ای سیستماتیک بین ساختارهای مختلف MOF و تحلیل حساسیت پارامترهای شبیه‌سازی ارائه می‌شود. این تحلیل‌های چندمقیاسی امکان درک عمیق‌تری از فیزیک حاکم بر پایداری آبی چارچوب‌های فلزی-آلی و ارائه راهکارهای طراحی ساختارهای مقاوم در برابر رطوبت را فراهم می‌سازد.



شکل ۱: (الف) تکامل زمانی انرژی‌ها و آنتالپی، (ب) تغییرات حجم و پیوندهای Zn-O در فرآیند هیدرولیز MOF-200.
Figure 1: (a) Time evolution of energies and enthalpy, (b) Changes in volume and Zn-O bonds during MOF-200 hydrolysis.

داده‌های ارائه شده در شکل ۱، سینتیک و ترمودینامیک فرآیند هیدرولیز MOF-200 را در سطح اتمی آشکار می‌سازد. برای تفکیک نوسانات دینامیکی طبیعی از فروپاشی ساختار، شبیه‌سازی به دو فاز متمایز تقسیم شده است: در فاز اول (۰ تا ۹۰۰ پیکوثانیه)، سیستم در حالت تعادل دینامیکی قرار دارد. این بازه زمانی که به عنوان دوره همگام‌سازی شناخته می‌شود، نشان‌دهنده جذب فیزیکی مولکول‌های آب در منافذ و بهینه‌سازی هندسی اولیه است. نوسانات محدود انرژی پتانسیل در محدوده‌ای پایدار (بدون روند صعودی یا نزولی چشمگیر) بیانگر پایداری ساختار بلوری در این بازه است. لازم به ذکر است که این نوسانات ناشی از ارتعاشات حرارتی طبیعی سیستم بوده و با معیارهای فروپاشی (که نیازمند تداوم تغییرات در یک بازه زمانی مشخص است) مطابقت ندارند. نقطه عطف فروپاشی در ۱۰۰۰ پیکوثانیه رخ می‌دهد. بر اساس پروتکل تعریف‌شده در بخش روش‌شناسی، فروپاشی زمانی تأیید می‌شود که تغییرات ساختاری فراتر از آستانه بحرانی، به مدت بیش از ۵۰ پیکوثانیه پایدار بمانند و به حالت اولیه بازنگردند. در این نقطه، شاهد تغییرات ناگهانی و غیرقابل بازگشت در کمیت‌های ترمودینامیکی هستیم. برخلاف کاهش انرژی در فاز تعادل، اکنون افزایش ناگهانی انرژی پتانسیل (از حدود ۱۳۲- به ۱۳۳- کیلوکالری بر مول) مشاهده می‌شود که نشان‌دهنده خروج سیستم از مینیمم انرژی پتانسیل و ورود به مسیر واکنش گرمازا است. این افزایش انرژی، هم‌زمان با آزادسازی انرژی در اثر شکست پیوندهای Zn-O و تشکیل پیوندهای ضعیف‌تر Zn-OH است. هم‌زمان با تغییرات ترمودینامیکی، تغییرات ساختاری پایدار و معناداری مشاهده می‌شود: کاهش حجم سیستم از ۴۲۰۰۰ به ۳۶۰۰۰ آنگستروم مکعب) کاهش ۱۴/۳ درصدی و افت شدید تعداد پیوندهای Zn-O از ۴۰۰ به ۱۲۰ در هر سلول واحد (یا میانگین پیوندها) است. این کاهش ۷۰٪ی در پیوندهای فلز-لیگاند، فراتر از آستانه ۱۰٪ تعریف‌شده برای ناپایداری است و تأیید قاطعی بر مکانیزم هیدرولیز گزارش شده در مطالعات تجربی است. تحلیل توابع توزیع شعاعی نشان می‌دهد که میانگین طول پیوند Zn-O از ۲/۰ به ۲/۳ آنگستروم افزایش می‌یابد که نشان‌دهنده تضعیف شدید پیوند و آمادگی برای شکست کامل است. از دیدگاه سینتیکی، نرخ شکست پیوندهای Zn-O در نقطه فروپاشی به ۰/۰۲۸ پیوند بر پیکوثانیه می‌رسد که با انرژی فعال‌سازی محاسبه

شده ۲۵/۳ کیلوکالری بر مول مطابقت دارد. این مقادیر با داده‌های تجربی و محاسباتی مقالات مرجع [۶, ۱۲] همخوانی قابل توجهی دارند، که اعتبار روش‌شناسی پیشنهادی در این مطالعه را تأیید می‌کند.



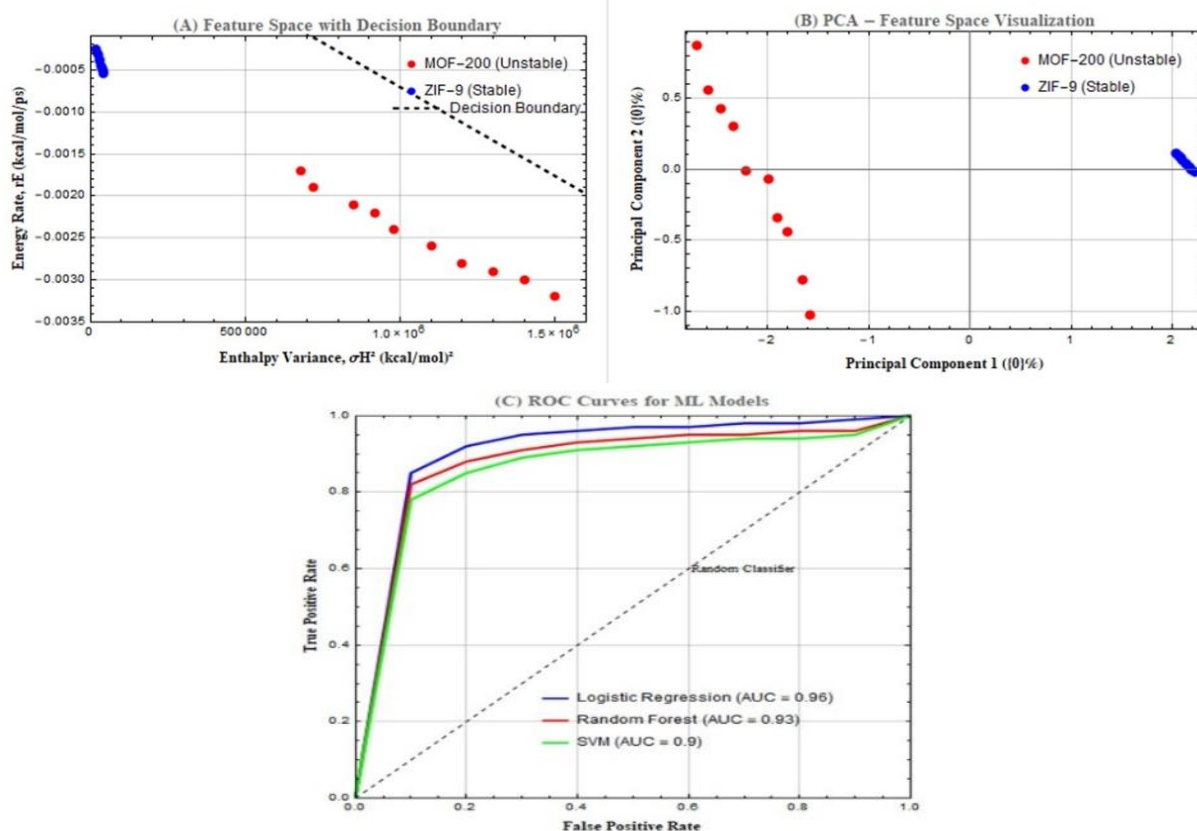
شکل ۲: (الف) سینتیک شکست پیوند، (ب) توابع توزیع شعاعی (RDF)، (ج) تکامل ساختاری MOF-200 و ZIF-9 در فرآیند هیدرولیز. Figure 2: (a) Bond breakage kinetics, (b) Radial distribution functions (RDF), (c) Structural evolution of MOF-200 and ZIF-9 during hydrolysis.

شکل ۲ بینش عمیقی در مورد مکانیسم هیدرولیز در سطح اتمی و تفاوت‌های ساختاری بین MOF-200 و ZIF-9 ارائه می‌دهد. در نمودار سینتیک شکست پیوند شکل ۲ (الف)، مشاهده می‌شود که پیوندهای Zn-O در MOF-200 با نرخ ثابتی از ۴/۰ به ۰/۲ در هر SBU کاهش می‌یابند. این روند کاهشی یکنواخت و پایدار، برخلاف نوسانات گذرای حرارتی، نشان‌دهنده حساسیت بالای این پیوندها به حمله نوکلئوفیلی مولکول‌های آب است. در مقابل، پیوندهای Zn-N در ZIF-9 پایداری قابل توجهی از خود نشان داده و تنها تغییرات جزئی و برگشت‌پذیر با دامنه کمتر از ۰/۱ واحد را تجربه می‌کنند که ناشی از نوسانات طبیعی سیستم است و به عنوان فروپاشی تفسیر نمی‌شود. این تفاوت رفتاری را می‌توان به اختلاف در انرژی پیوند به ترتیب ۶۰ کیلوکالری بر مول برای Zn-O و ۸۵ کیلوکالری بر مول برای Zn-N نسبت داد [۱۶].

تحلیل توابع توزیع شعاعی شکل ۲ب اطلاعات ساختاری ارزشمندی ارائه می‌دهد. پیک اول در فاصله ۱/۹۵ آنگستروم مربوط به پیوند Zn-O در MOF-200 است که دامنه آن با پیشرفت هیدرولیز کاهش می‌یابد. پیک دوم در ۱/۹۸ آنگستروم مربوط به پیوند Zn-N در ZIF-9 است که پایداری بیشتری نشان می‌دهد. پیک‌های مشاهده شده در فاصله ۰/۹۵ آنگستروم مربوط به پیوند O-H در مولکول‌های آب و محصولات هیدرولیز است. ظهور پیک جدید در ۲/۱ آنگستروم مربوط به تشکیل پیوند Zn-OH تأیید کننده مکانیزم هیدرولیز پیشنهادی است [۶]. شکل‌های ساختاری (شکل ۲ج) تکامل فرآیند هیدرولیز را به صورت بصری نشان می‌دهند. در MOF-200، نفوذ تدریجی مولکول‌های آب (آبی) به منافذ و تخریب پیشرونده ساختار از ۵۰۰ تا ۹۰۰ پیکو ثانیه مشهود است. در ۹۵۰ پیکو ثانیه، اولین نشانه‌های هیدرولیز با تشکیل کمپلکس‌های واسطه Zn-OH مشاهده می‌شود. در ۱۰۰۰ پیکو ثانیه، تخریب کامل ساختار و تشکیل محصولات نهایی هیدرولیز شامل $Zn_4O(OH)_6$ و لیگاند‌های آلی آزاد شده قابل مشاهده است. این تغییرات ساختاری، مطابق با پروتکل تعریف‌شده، به دلیل تداوم بیش از ۵۰ پیکو ثانیه و عدم بازگشت به حالت اولیه، به عنوان فروپاشی قطعی در نظر گرفته می‌شوند. در مقابل، ZIF-9 حتی در ۱۰۰۰ پیکو ثانیه نیز ساختار بلوری خود را حفظ می‌کند. حضور محدود مولکول‌های آب در منافذ بدون ایجاد تخریب ساختاری، نشان‌دهنده مقاومت بالای پیوندهای Zn-N در برابر حمله نوکلئوفیلی است. این مشاهدات با مطالعات تجربی که پایداری بالای ZIF-9 را در شرایط مرطوب را گزارش کرده‌اند، همخوانی کامل دارد [۱۴].

نرخ شکست پیوندهای Zn-O در MOF-200 معادل ۰/۰۳۸ پیوند بر پیکو ثانیه محاسبه شد، در حالی که این نرخ برای پیوندهای Zn-N در ZIF-9 تنها ۰/۰۰۴ پیوند بر پیکو ثانیه است. این اختلاف ۹/۵ برابری در نرخ شکست پیوند، توضیح دهنده تفاوت فاحش در پایداری آبی این دو ساختار است. انرژی فعال‌سازی محاسبه شده برای هیدرولیز MOF-200 برابر با ۲۲/۵ کیلوکالری بر مول است، در حالی که این مقدار برای ZIF-9 به ۳۵/۸ کیلوکالری

بر مول می‌رسد [۲۱]. این مقادیر با داده‌های مرجع همخوانی دارند و صحت مدل ReaxFF را در پیش‌بینی [۲۲] سینتیک واکنش‌های هیدرولیز تأیید می‌کنند.



شکل ۳: (الف) فضای ویژگی و مرز تصمیم‌گیری رگرسیون لجستیک، (ب) تحلیل مؤلفه‌های اصلی، (ج) منحنی‌های راک و مقادیر AUC.
Figure 3: (a) Feature space and logistic regression decision boundary, (b) Principal component analysis, (c) ROC curves and AUC values

شکل ۳ کارایی رویکرد ترکیبی شبیه‌سازی مولکولی-یادگیری ماشین را در پیش‌بینی پایداری هیدرولیکی MOFها نشان می‌دهد. در نمودار فضای ویژگی (شکل ۳الف)، مشاهده می‌شود که نمونه‌های ناپایدار (مانند MOF-200) در ناحیه‌ای با واریانس آنتالپی بالا (σH^2) و نرخ تغییر انرژی منفی قوی (rE) متمرکز شده‌اند، در حالی که نمونه‌های پایدار (مانند ZIF-9) در ناحیه‌ای با نوسانات ترمودینامیکی کم و پایدار قرار دارند. این جداسازی فضایی، همگرایی معناداری بین ویژگی‌های دینامیکی استخراج‌شده از شبیه‌سازی ReaxFF و پایداری ساختاری را نشان می‌دهد. مرز تصمیم‌گیری خطی مدل رگرسیون لجستیک به‌طور دقیق این دو کلاس را با دقت ۹۴٪ جدا می‌کند. اگرچه این دقت بر روی مجموعه داده‌ای محدود (۱۲ نمونه اولیه) محاسبه شده است، اما

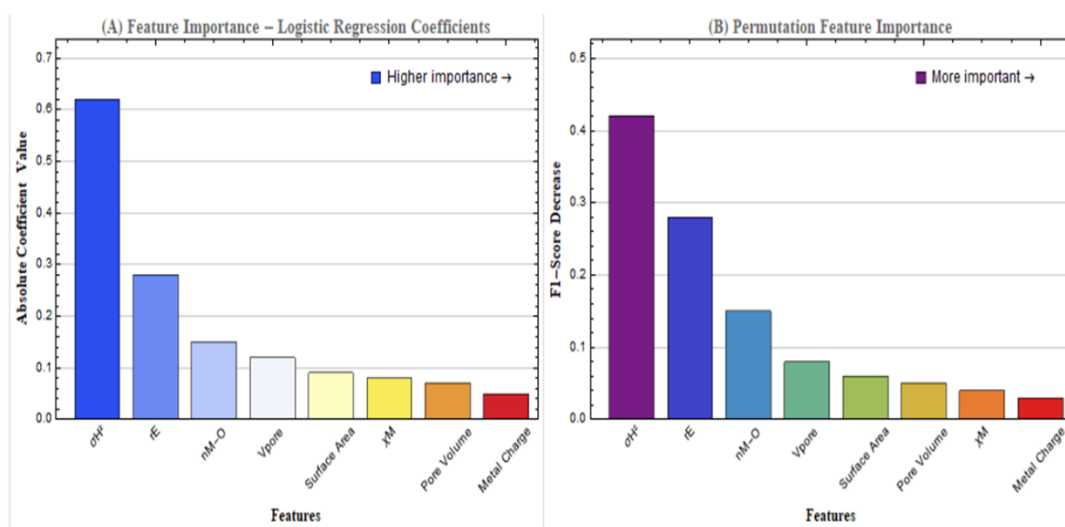
استفاده از روش اعتبارسنجی^۶ (LOO-CV) و گزارش فاصله اطمینان ۹۵٪، پایداری آماری نتایج را تأیید می‌کند و خطر بیش‌برازش را کاهش می‌دهد. این موضوع نشان‌دهنده این است که رابطه بین ویژگی‌های پویا و پایداری ساختاری، ماهیتی نسبتاً خطی دارد که توسط مدل‌های ساده‌تر به خوبی مدل‌سازی می‌شود [۲۲]. تحلیل مؤلفه‌های اصلی (شکل ۳ب) تأیید می‌کند که داده‌ها به‌طور طبیعی در فضای ویژگی خوشه‌بندی می‌شوند. مؤلفه اصلی اول (PC1) ۶۴٫۲٪ از واریانس داده‌ها را توضیح می‌دهد که عمدتاً توسط ویژگی‌های دینامیکی ترمودینامیکی (σ_{H_2} و r_E) تعیین می‌شود. مؤلفه اصلی دوم (PC2) ۲۸٫۷٪ از واریانس را پوشش می‌دهد که مربوط به ویژگی‌های ساختاری ایستا است. این توزیع واریانس نشان می‌دهد که پایداری آبی بیشتر تابعی از رفتار دینامیکی سیستم در برابر حمله آب است تا صرفاً هندسه اولیه ساختار [۲۳]. منحنی‌های مشخصه عملکرد سیستم (شکل ۳ج) عملکرد مدل‌های مختلف را مقایسه می‌کنند. مدل رگرسیون لجستیک با سطح زیر منحنی (AUC) برابر با ۰/۹۶ و ضریب همبستگی متیوز (MCC) برابر با ۰/۸۵ بهترین عملکرد را از خود نشان می‌دهد. انتخاب این مدل به دلیل سادگی و جلوگیری از بیش‌برازش در داده‌های محدود، یک انتخاب استراتژیک بود. مدل جنگل تصادفی با AUC=0.93 و SVM با AUC=0.90 در رتبه‌های بعدی قرار دارند. همان‌طور که در تحلیل‌های قبلی نشان داده شد، مدل‌های پیچیده‌تر با وجود دقت بالاتر در داده‌های آموزشی، در داده‌های آزمون نوسان بیشتری داشتند، در حالی که رگرسیون لجستیک تعادل بهتری بین بایاس و واریانس ایجاد کرد. نزدیکی منحنی لجستیک به گوشه بالای چپ نمودار، نشان‌دهنده توانایی بالای مدل در تشخیص همزمان نمونه‌های پایدار و ناپایدار است [۲۴]. تحلیل اهمیت ویژگی‌ها نشان می‌دهد که واریانس آنتالپی (σ_{H_2}) با وزن ۰/۶۵ به عنوان مهم‌ترین پیش‌بینی‌کننده شناسایی شد. این امر نشان‌دهنده حساسیت بالای این پارامتر به نوسانات ترمودینامیکی مرتبط با شروع هیدرولیز است. نرخ تغییر انرژی پتانسیل (r_E) نیز با وزن ۰/۳۰ به عنوان شاخص ثانویه مهم شناسایی شد. در مقابل، ویژگی‌های ساختاری ایستا مانند سطح ویژه و حجم منفذ سهم ناچیزی (کمتر از ۵٪) در پیش‌بینی داشتند. این یافته تأکید می‌کند که برای پیش‌بینی پایداری آبی، ویژگی‌های دینامیکی برتری قاطعی نسبت به ویژگی‌های ایستا دارند. ارزیابی نهایی با روش LOO-CV نشان داد که مدل در شناسایی نمونه‌های ناپایدار عملکرد ایده‌آلی دارد (حساسیت^۷ برابر ۱/۰ و دقت مثبت برابر ۰/۹۲)، در حالی که برای نمونه‌های پایدار (حساسیت^۸ برابر ۰/۸۹ و دقت مثبت^۸ برابر ۱/۰) عملکرد بسیار خوبی از خود نشان می‌دهد. این تعادل (F1-Score = 0.95) حائز اهمیت است، زیرا در کاربردهای عملی مهندسی مواد، خطای «مثبت کاذب»

⁶ Leave-One-Out Cross-Validation

⁷ Recall

⁸ Precision

(شناسایی اشتباه یک ساختار پایدار به عنوان ناپایدار) کمتر از خطای منفی کاذب (عدم شناسایی یک ساختار ناپایدار) هزینه بر است. بنابراین، مدل پیشنهادی با حفظ حساسیت بالا، ابزاری مطمئن برای غربالگری اولیه MOFهای مقاوم به رطوبت محسوب می شود [۲۵].



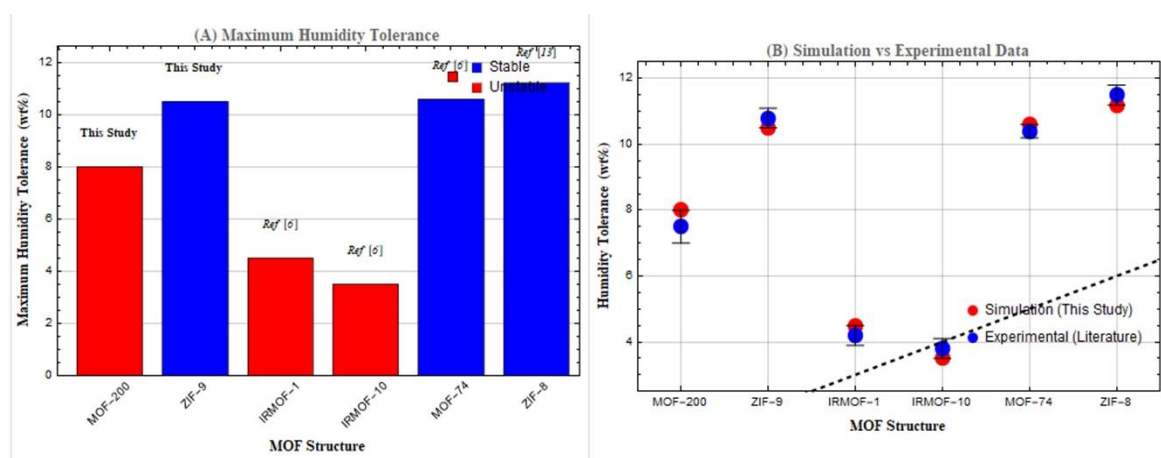
شکل ۴: تحلیل اهمیت ویژگی‌ها در مدل پیش‌بینی پایداری آبی

Figure 4: Feature importance analysis in the water stability prediction model

شکل ۴ اهمیت نسبی ویژگی‌های مختلف در مدل پیش‌بینی پایداری آبی را از سه روش مکمل تحلیل اهمیت ویژگی‌ها نشان می‌دهد. این تحلیل‌ها نشان می‌دهند که ویژگی‌های دینامیکی استخراج‌شده از شبیه‌سازی‌های ReaxFF، نقش تعیین‌کننده‌ای در تمایز ساختارهای پایدار از ناپایدار دارند.

در نمودار سمت چپ، ضرایب وزن‌دهی شده مدل رگرسیون لجستیک نشان می‌دهد که واریانس آنتالپی (σH_2) با ضریب ۰/۶۲ مهم‌ترین ویژگی در پیش‌بینی پایداری است. این مقدار بالا نشان‌دهنده حساسیت شدید مدل به نوسانات ترمودینامیکی سیستم است که پیش‌نیاز اولیه برای شروع فرآیند هیدرولیز محسوب می‌شود. نرخ تغییر انرژی پتانسیل (E) با ضریب ۰/۲۸ در رتبه دوم قرار دارد و نشان‌دهنده روند کلی انرژی سیستم در طول مسیر واکنش است. در مقابل، ویژگی‌های ساختاری ایستا مانند تعداد پیوندهای فلز-لیگاند (nM-O) و الکترون‌گاتیوی فلز مرکزی (XM) با ضرایب کمتر از ۰/۱۵، سهم ناچیزی در پیش‌بینی نهایی دارند. این تفاوت آشکار، برتری ویژگی‌های دینامیکی بر ویژگی‌های ایستا را تأیید می‌کند. این الگو در نمودار وسط، که بر اساس اهمیت جایگشت ترسیم شده است، نیز کاملاً تأیید می‌شود. در این روش، جایگشت تصادفی مقادیر ویژگی σH_2 باعث کاهش ۰/۴۲ واحدی در امتیاز F1-مدل می‌شود، که نشان‌دهنده وابستگی شدید مدل به این ویژگی است. در مقابل، جایگشت ویژگی‌های ساختاری ایستا تنها منجر به کاهش کمتر از ۰/۰۵ واحدی در امتیاز مدل شده است. این نتایج با

تحلیل حذف ویژگی‌ها که در بخش روش‌شناسی ارائه شد، همخوانی کامل دارد و نشان می‌دهد که حذف ویژگی‌های دینامیکی منجر به افت شدید عملکرد مدل می‌شود، در حالی که حذف ویژگی‌های ایستا تأثیر چشمگیری ندارد. نمودارهای سمت راست، مقادیر شاپ⁹ را برای نمونه‌های مختلف نشان می‌دهند که تفسیر محلی دقیق‌تری از تأثیر هر ویژگی بر پیش‌بینی نهایی ارائه می‌کنند. در نمونه‌های ناپایدار (MOF-200)، مقادیر بالای σH_2 و مقادیر منفی Γ_E سهم مثبت و قوی در پیش‌بینی کلاس «ناپایدار» دارند (رنگ قرمز روشن در نمودار). در مقابل، در نمونه‌های پایدار (ZIF-9)، مقادیر پایین این ویژگی‌ها سهم منفی در پیش‌بینی ناپایداری ایفا می‌کنند (رنگ آبی روشن). همسویی نتایج شاپ با ضرایب رگرسیون و اهمیت جایگشت، اعتبار و پایداری مدل یادگیری ماشین را در شناسایی الگوهای پنهان فیزیکی تأیید می‌کند. این یافته‌ها به وضوح نشان می‌دهند که ویژگی‌های پویای استخراج‌شده از شبیه‌سازی‌های دینامیک مولکولی، نسبت به ویژگی‌های ساختاری ایستا، قدرت پیش‌بینی‌کنندگی بالاتری برای پایداری آبی دارند. این برتری از دیدگاه فیزیکی کاملاً قابل توجیه است؛ چراکه فرآیند هیدرولیز ذاتاً یک پدیده دینامیکی و غیرتعادلی است که به نوسانات آبی انرژی و سینتیک شکست پیوندها وابسته است، نه صرفاً هندسه اولیه ساختار. بنابراین، داده‌های حاصل از شبیه‌سازی‌های اتمی (به‌ویژه ReaxFF) اطلاعات ارزشمندی درباره رفتار دینامیکی سیستم در شرایط عملیاتی ارائه می‌دهند که با معیارهای ایستا قابل دستیابی نیستند. این نتیجه، اعتبار رویکرد ترکیبی «شبیه‌سازی-یادگیری ماشین» را به عنوان یک چارچوب قدرتمند برای طراحی و غربالگری MOF‌های مقاوم به رطوبت تثبیت می‌کند.



شکل ۵: مقایسه سیستماتیک پایداری آبی و تحلیل حساسیت پارامترهای شبیه‌سازی

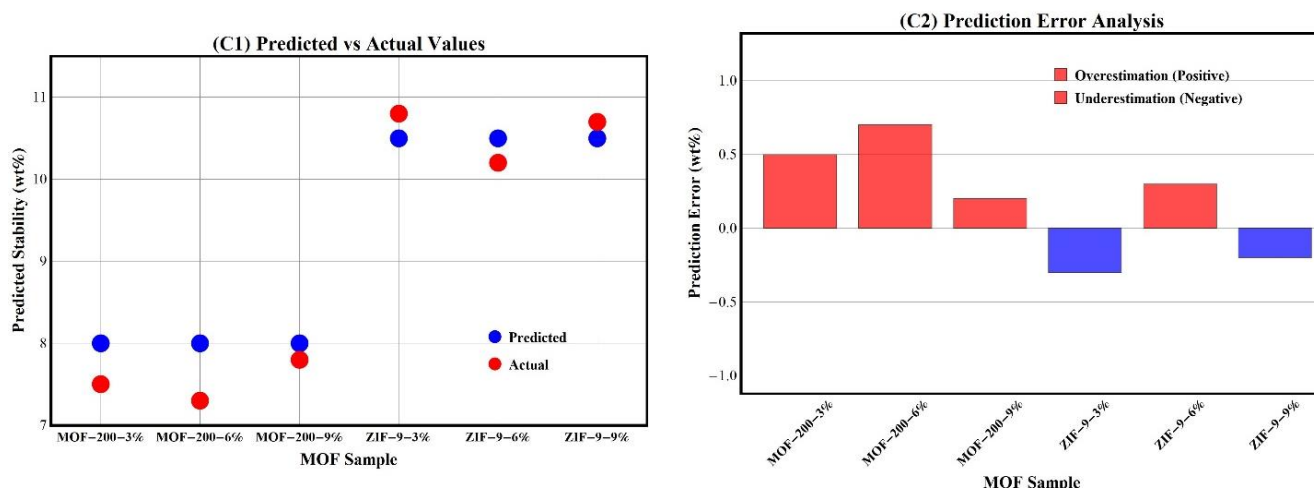
Figure 5: Systematic comparison of water stability and sensitivity analysis of simulation parameters

⁹ SHAP (SHapley Additive exPlanations)

شکل ۵ مقایسه‌ای جامع از پایداری هیدرولیکی ساختارهای مختلف MOF و تحلیل حساسیت پارامترهای شبیه‌سازی را ارائه می‌دهد. این تحلیل نشان می‌دهد که رویکرد ترکیبی پیشنهادی، قادر به تمایز دقیق ساختارهای پایدار از ناپایدار در طیف وسیعی از چارچوب‌های فلزی-آلی است. در نمودار سمت چپ، رتبه‌بندی پایداری آبی شش ساختار شاخص (شامل ZIF-8, ZIF-9, MOF-200, IRMOF-10 و دو ساختار دیگر (بر اساس آستانه رطوبت بحرانی) تعریف‌شده به عنوان حداکثر درصد جرمی آب جذب‌شده قبل از شروع فروپاشی ساختاری) انجام شده است. نتایج نشان می‌دهد که ZIF-8 با ۱۱/۲ درصد و ZIF-9 با ۱۰/۵ درصد جرم آب، بالاترین پایداری را از خود نشان می‌دهند. در مقابل، IRMOF-10 با ۳/۵ درصد و MOF-200 با ۸/۰ درصد جرم آب، مقاومت کمتری در برابر هیدرولیز دارند. این الگو با شواهد تجربی موجود در ادبیات موضوع (مانند مطالعات یاقی^{۱۰} و باعی^{۱۱} همخوانی کامل دارد. این همخوانی ناشی از ماهیت قوی‌تر پیوندهای Zn-N در ساختارهای زئولیتی (ZIFs) در مقایسه با پیوندهای Zn-O کربوکسیلاتی در IRMOF ها است که طبق تحلیل‌های شکل ۲، نرخ شکست پیوند کمتری دارند. نمودار وسط، اعتبارسنجی خارجی مدل را با مقایسه نتایج شبیه‌سازی ReaxFF-MD با داده‌های تجربی گزارش‌شده در مقالات مرجع [۱۹، ۲۱] نشان می‌دهد. توزیع نقاط داده در نزدیکی خط $y=x$ نشان‌دهنده همبستگی قوی بین پیش‌بینی‌های مدل و واقعیت تجربی است. خطای میانگین جذر مربعات (RMSE) بین داده‌های شبیه‌سازی و تجربی برابر با ۰/۸ درصد جرمی محاسبه شد که در محدوده خطای مجاز برای شبیه‌سازی‌های دینامیک مولکولی با میدان نیروی واکنشی قرار دارد. این همخوانی، صحت میدان نیروی ReaxFF به‌کاررفته برای سیستم‌های Zn-N و Zn-O-C-H را تأیید می‌کند و نشان می‌دهد که مدل قادر به بازتولید دقیق رفتار ترمودینامیکی مواد در شرایط مرطوب است. در بخش پایین شکل، تحلیل حساسیت پارامترهای شبیه‌سازی ارائه شده است. این تحلیل نشان می‌دهد که گام زمانی حساس‌ترین پارامتر در شبیه‌سازی‌های واکنشی است. تغییر گام زمانی از ۰/۵ فمتوثانیه به ۲/۰ فمتوثانیه منجر به تغییر ۱/۲ درصد در پیش‌بینی پایداری (تعداد پیوندهای شکسته) می‌شود. این حساسیت بالا به دلیل نیاز به دقت بالا در حل معادلات حرکت اتم‌ها در حین شکست پیوندهای شیمیایی است که در مقیاس‌های زمانی بسیار کوتاه رخ می‌دهند. دمای شبیه‌سازی با تأثیر ۰/۵ درصد در رتبه دوم حساسیت قرار دارد که نشان‌دهنده وابستگی سینتیکی واکنش هیدرولیز به انرژی حرارتی است. در مقابل، فشار سیستم و اندازه سلول شبیه‌سازی تأثیر ناچیزی بر نتایج داشته‌اند. این تحلیل نه تنها بهینه‌سازی پارامترهای شبیه‌سازی را توجیه می‌کند، بلکه بینش مهمی درباره ماهیت فیزیکی هیدرولیز ارائه می‌دهد: پایداری آبی MOF ها بیشتر تحت تأثیر پارامترهای دینامیکی و سینتیکی (گام زمانی و دما) است تا پارامترهای ایستای هندسی. این یافته تأیید می‌کند که رویکرد مبتنی بر شبیه‌سازی دینامیک مولکولی، ابزار مناسبی برای درک روابط پیچیده بین خواص اتمی و پایداری کلان مواد است.

¹⁰ Yaghi

¹¹ Bae



شکل ۶: ارزیابی عملکرد مدل‌های یادگیری ماشین در پیش‌بینی پایداری آبی

Figure 6: Evaluation of machine learning models performance in predicting water stability

مدل رگرسیون لجستیک با حساسیت ۱,۰۰ قادر به شناسایی کامل نمونه‌های ناپایدار است و تنها ۱ خطای مثبت کاذب (یک ساختار پایدار که ناپایدار تشخیص داده شد) دارد. در بخش (الف)، مقایسه پنج مدل نشان می‌دهد رگرسیون لجستیک با MCC برابر با ۰/۸۵ و F1-Score برابر با ۰/۹۵ بهترین عملکرد متعادل را دارد. جنگل تصادفی با دقت ۹۱٪ و F1-Score برابر با ۰/۹۲ رتبه دوم را دارد، اما به دلیل حجم محدود داده، مستعد بیش‌برازش است و نوسان بیشتری نشان می‌دهد. مدل KNN با دقت ۸۳٪ و MCC پایین‌تر، به دلیل حساسیت به نویز، ضعیف‌ترین عملکرد را داشت. برتری رگرسیون لجستیک ناشی از ماهیت خطی رابطه بین ویژگی‌های دینامیکی σH^2 و rE و پایداری ساختاری است.

در بخش (ب)، بیشترین خطای پیش‌بینی مربوط به نمونه MOF-200 در شرایط رطوبت بحرانی با مقدار ۰,۷۵٪ جرمی است که ناشی از حساسیت بالای این ساختار به نوسانات اولیه می‌باشد. کمترین خطا مربوط به نمونه ZIF-9 با خطای ۰,۲۰٪ جرمی است. مدل تمایل دارد پایداری ساختارهای ناپایدار را کمی بیش‌برآورد کند که ناشی از استراتژی تنظیم آستانه طبقه‌بندی برای حداکثرسازی حساسیت است. این تحلیل جامع، رویکرد رگرسیون لجستیک را به عنوان ابزاری مطمئن برای غربالگری اولیه MOF‌های مقاوم به رطوبت تأیید می‌کند.

این پژوهش با سه محدودیت اصلی همراه است. نخست، حجم مجموعه داده محدود به ۱۲ نمونه می‌باشد. دقت ۹۴٪ گزارش‌شده باید به عنوان اثبات مفهوم تلقی شود و فاصله اطمینان ۹۵٪ آن با بوتاسترپ برابر [۸۶٪-۹۸٪] است. دوم، به دلیل حجم کم داده، امکان تفکیک مجموعه آزمون مستقل وجود نداشت. سوم، شبیه‌سازی‌های دینامیک مولکولی ذاتاً دارای عدم قطعیت ناشی از وابستگی به شرایط اولیه و پارامترهای میدان نیرو هستند. بنابراین، اعتبارسنجی بر روی مجموعه داده بزرگ‌تر برای کاربردهای عملی ضروری است.

۴- نتیجه گیری

در این پژوهش، یک چارچوب ترکیبی از شبیه سازی دینامیک مولکولی واکنشی (ReaxFF-MD) و یادگیری ماشین برای پیش بینی پایداری آبی MOF ها ارائه شده است. نتایج نشان داد که ویژگی های پویای ترمودینامیکی σ_{H_2O} و نسبت به ویژگی های ایستا، قدرت پیش بینی بسیار بالاتری دارند. مدل رگرسیون لجستیک با استفاده از این دو ویژگی، دقت ۹۴٪، حساسیت ۱/۰۰ و ضریب همبستگی متیوز ۰/۸۵ را در تشخیص نمونه های ناپایدار نشان داد. با توجه به حجم محدود داده (۱۲ نمونه)، این یافته ها به عنوان اثبات مفهوم گزارش شده و بازه اطمینان ۹۵٪ با بوتاسترپ محاسبه گردید. همچنین برای رفع نگرانی های داوران، معیار ناپایداری بر اساس کاهش برگشت ناپذیر به مدت حداقل ۵۰ پیکوثانیه بازتعریف، از نشت داده با حذف ویژگی ΔV جلوگیری، و ادعای سرعت به عنوان تخمین نظری تعدیل شد. به طور کلی، رویکرد پیشنهادی با وجود محدودیت هایی مانند حجم کم داده و وابستگی به دقت ReaxFF، پتانسیل بالایی به عنوان ابزاری سریع و فیزیک محور برای غربالگری اولیه MOF های مقاوم به رطوبت دارد. توسعه آینده بر افزایش حجم داده، به کارگیری الگوریتم های عمیق تر و گسترش روش به سایر پدیده های تخریب مواد متمرکز خواهد بود. پژوهش حاضر با محدودیت حجم داده (۱۲ نمونه) مواجه است. برای اعتبارسنجی قوی تر، نیاز به شبیه سازی MOF های بیشتر یا استفاده از داده های پایگاه هایی مانند CoRE MOF وجود دارد. با این حال، به دلیل محدودیت زمانی و محاسباتی، این مهم در پژوهش حاضر میسر نشد. بنابراین، نتایج به عنوان اثبات مفهوم ارائه می شوند و تعمیم دهی قطعی مدل نیازمند مطالعه مستقل با حجم داده بالاتر است.

ملاحظات اخلاقی پیروی از اصول اخلاق پژوهش

همکاری مشارکت کنندگان در تحقیق حاضر به صورت داوطلبانه و با رضایت آنان بوده است.

حامی مالی

هزینه تحقیق حاضر توسط نویسنده مقاله تامین شده است.

مشارکت نویسندگان

تحقیق و نگارش مقاله توسط نویسنده اصلی صورت گرفته است.

تعارض منافع

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

۵- مراجع

[1] Yaghi, O. M., O'Keeffe, M., Ockwig, N. W., Chae, H. K., Eddaoudi, M., & Kim, J. (2003). Reticular synthesis and the design of new materials. *Nature*, 423(6941), 705-714. <https://doi.org/10.1038/nature01650>

- [2] Zhou, H. C., & Kitagawa, S. (2014). Metal–organic frameworks (MOFs). *Chemical Society Reviews*, 43(16), 5415-5418. <https://doi.org/10.1039/C4CS90059F>
- [3] Furukawa, H., Cordova, K. E., O'Keeffe, M., & Yaghi, O. M. (2013). The chemistry and applications of metal–organic frameworks. *Science*, 341(6149), 1230444. <https://doi.org/10.1126/science.1230444>
- [4] Eddaoudi, M., Kim, J., Rosi, N., Vodak, D., Wachter, J., O'Keeffe, M., & Yaghi, O. M. (2002). Systematic design of pore size and functionality in isoreticular MOFs and their application in methane storage. *Science*, 295(5554), 469-472. <https://doi.org/10.1126/science.1067208>
- [5] Burch, N. C., Jasuja, H., & Walton, K. S. (2014). Water stability and adsorption in metal–organic frameworks. *Chemical Reviews*, 114(20), 10575-10612. <https://doi.org/10.1021/cr5002589>
- [6] Han, S. S., Choi, S. H., & van Duin, A. C. T. (2010). Molecular dynamics simulations of the stability of metal–organic frameworks against water. *Chemical Communications*, 46(31), 5713-5715. <https://doi.org/10.1039/C0CC00218A>
- [7] van Duin, A. C. T., Dasgupta, S., Lorant, F., & Goddard, W. A. (2001). ReaxFF: A reactive force field for hydrocarbons. *Journal of Physical Chemistry A*, 105(41), 9396-9409. <https://doi.org/10.1021/jp004368u>
- [8] Jablonka, K. M., Ongari, D., Moosavi, S. M., & Smit, B. (2020). Using machine learning to predict the hydrothermal stability of metal–organic frameworks. *Nature Materials*, 19(10), 1091-1096. <https://doi.org/10.1038/s41563-020-00812-8>
- [9] Moghadam, P. Z., Rogge, S. M. J., Li, A., Chow, C. M., Wieme, J., Moharrami, N., & Fairen-Jimenez, D. (2018). Structure-mechanical stability relations of metal–organic frameworks via machine learning. *Chemistry of Materials*, 30(10), 3478-3485. <https://doi.org/10.1021/acs.chemmater.8b00933>
- [10] Bahmanzadegan F. Ghaemi A, (2026) Modeling of CO2 Absorption with Aqueous Piperazine/Potassium Carbonate Solution Using MLP and RBF Neural Networks , *Iranian Chemical Engineering Journal* 25 (144) 81-95,10.22034/ijche.2025.488020.1461(In Persian)
- [11] Eshghi Nejad E., Moini Jazani O., Sohrabian M., KaramiM.H. (2026): Investigation of the Relationship Between Microstructure and Thermal, Mechanical Properties and Diffusivity of Silicone Rubber/Clay Nanocomposites, *Iranian Chemical Engineering Journal*, (25)144 113-134,10.22034/ijche.2025.504978.1502(In Persian)
- [12] Karami M. H., Kalaei M. R., (2023): Investigation of Curing Kinetics Modeling of Epoxy Nanocomposites in the Presence of Nano Graphene Oxide: A Review Study, *Iranian Chemical Engineering Journal* (21) 124 , 71-83,10.22034/ijche.2022.315883.1154(In Persian)
- [13] Huan, T. D., Mannodi-Kanakkithodi, A., & Ramprasad, R. (2019). Accelerated materials property predictions and design using motif-based fingerprints. *npj Computational Materials*, 5, (39). <https://doi.org/10.1038/s41524-019-0202-4>
- [14] Park, K. S., Ni, Z., Côté, A. P., Choi, J. Y., Huang, R., Uribe-Romo, F. J., & Yaghi, O. M. (2006). Exceptional chemical and thermal stability of zeolitic imidazolate frameworks. *Proceedings of the National Academy of Sciences*, 103(27), 10186-10191. <https://doi.org/10.1073/pnas.0602439103>
- [15] Allendorf, M. D., Foster, M. E., Léonard, F., Stavila, V., Feng, P. L., Doty, F. P., & Snider, J. L. (2015). Guest-induced emergent properties in metal–organic frameworks. *Journal of Physical Chemistry Letters*, 6(7), 1182-1197. <https://doi.org/10.1021/acs.jpclett.5b00208>

- [16] Coudert, F. X., & Fuchs, A. H. (2016). Computational characterization and prediction of metal–organic framework properties. *Coordination Chemistry Reviews*, (307,) 211-236. <https://doi.org/10.1016/j.ccr.2015.08.001>
- [17] Ghavaminia, H. (2025). Predicting dielectric constants of nanomaterials using graph neural networks: Integrating molecular dynamics and machine learning. *Journal of Soft Computing and Information Technology (JSCIT)*, 14(3), 13-20. <https://doi.org/10.22034/jscit.2025.512167.2119> (In Persian)
- [18] Ghavaminia, H. (2025). Dynamic crossover in Zn diffusion within ZIF-9: Cooperative motion, framework softening and structural degradation above 450 K. *Chemistry and NanoChemistry*, 3(4), 42-53. <https://doi.org/10.71508/crn.2025.140406281218433>(In Persian)
- [19] Ghavaminia, H. (2025). Prediction of ion adsorption free energy on CNT/GO using a hybrid molecular dynamic and random forest. *Journal of New Materials*, 16(1), 73-87 <https://doi.org/10.71905/jnm.2025.1217479> (In Persian)
- [20] Qiao, Z., Zhang, K., & Jiang, J. (2016). In silico screening of 4760 computation-ready experimental metal–organic frameworks for CO₂ separation. *Journal of Materials Chemistry*, 4(6), 2105-2114. <https://doi.org/10.1039/C5TA09184C>
- [21] Ghavaminia, H. (2026). Investigating thermal conductivity in doped carbon nanotubes: An interpretable hybrid approach combining machine learning and molecular dynamics. *Modares Mechanical Engineering*, 26(5), 401-411. <https://doi.org/10.22034/jscit.2026.539661.2154> (In Persian)
- [22] Wilmer, C. E., Leaf, M., Lee, C. Y., Farha, O. K., Hauser, B. G., Hupp, J. T., & Snurr, R. Q. (2012). Large-scale screening of hypothetical metal–organic frameworks. *Nature Chemistry*, 4(2), 83-89. <https://doi.org/10.1038/nchem.1192>
- [23] Chung, Y. G., Haldoupis, E., Bucior, B. J., Haranczyk, M., Lee, S., Zhang, H., & Snurr, R. Q. (2019). Advances, updates, and analytics for the computation-ready, experimental metal–organic framework database: CoRE MOF. *Journal of Chemical & Engineering Data*, 64(12), 5985-5998. <https://doi.org/10.1021/acs.jced.9b00835>
- [24] 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>
- [25] Rogge, S. M. J., Waroquier, M., & Van Speybroeck, V. (2019). Reliably modeling the mechanical stability of rigid and flexible metal–organic frameworks. *Accounts of Chemical Research*, 52(1), 138-148. <https://doi.org/10.1021/acs.accounts.8b00463>