Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1390−1399
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8329

Stacking Machine Learning Approach for Predicting Thermal Stability of Zinc–Metal Organic Frameworks (Zn-MOF)

Ananta Surya Pratama[1], Muhammad Diva Irnanda[1], Taufiqul Umam[1], Dandy Prasetyo Nugroho[1], Harun Al Azies[2,*]

[1] Faculty of Computer Science, Informatics Engineering, Universitas Dian Nuswantoro, Semarang, Indonesia
[2] Faculty of Computer Science, Research Center for Quantum Computing and Materials Informatics, Universitas Dian Nuswantoro, Semarang, Indonesia
Email: [1]111202315107@mhs.dinus.ac.id, [2]111202315484@mhs.dinus.ac.id, [3]111202315273@mhs.dinus.ac.id, [4]111202415645@mhs.dinus.ac.id, [5,*]harun.alazies@dsn.dinus.ac.id
Email Penulis Korespondensi: harun.alazies@dsn.dinus.ac.id
Submitted: **04/09/2025**; Accepted: **29/09/2025**; Published**: 29/09/2025**

Abstract−Thermal stability is a fundamental parameter that determines the feasibility of Metal Organic Frameworks (MOF) for high-temperature industrial applications, including catalysis, gas purification, and energy storage. Experimental evaluation of thermal stability, while accurate, is often costly and time-consuming, highlighting the need for computational prediction models that are both efficient and dependable. This study develops a Quantitative Structure Property Relationship (QSPR) model using a stacking ensemble regression framework to predict the thermal stability of Zn-MOFs. The stacking approach combines Linear Regression, Lasso Regression, and Huber Regression as base learners, with Linear Regression serving as the meta-model, thereby leveraging the complementary strengths of individual algorithms. Results demonstrate that the stacking ensemble consistently outperformed all single models, delivering highly reliable predictions that remained stable across multiple validation scenarios. Furthermore, external validation with experimental data confirmed the model's robustness and its ability to generalize beyond the training dataset. These findings underline the reliability of stacking as not only a tool for improving accuracy but also for ensuring predictive stability and reproducibility. The study highlights the potential of machine learning, particularly ensemble methods, as a powerful and trustworthy predictive framework for the rational design of thermally stable MOFs, offering both scientific and industrial significance in sustainable energy applications.

Keywords: Ensemble Learning; Machine Learning; Metal Organic Framework; Stacking; Thermal Stability

1. INTRODUCTION

The accelerated pace of climate change has placed the world in an environmental emergency, demanding a transformation toward low-emission energy systems. According to the International Energy Agency (IEA), global CO_2 emissions from energy and industrial processes reached a record 37.8 gigatonnes in 2024, underscoring the urgent need to advance clean energy technologies capable of mitigating greenhouse gas emissions[1]. The net-zero emission targets set by many countries, including Indonesia, further highlight the demand for functional materials that can support sustainable energy technologies and reduce dependence on fossil fuels.At the industrial scale, numerous strategic technologies such as heterogeneous catalysis, gas purification, energy storage, and sustainable chemical processes operate under high-temperature conditions to maintain structural integrity and functional performance. This makes the thermal stability of materials a critical factor determining the success of these applications. Unfortunately, the discovery and development of high-resilience materials remain constrained by experimental methods that are time-consuming, costly, and labor-intensive, limiting the pace of material innovation. This situation creates a pressing need for intelligent computational approaches that can predict material properties quickly, accurately, and efficiently [2].

The emergence of Metal–Organic Frameworks (MOF) has garnered widespread attention due to their high porosity, chemical flexibility, and large specific surface area [3]. The modular structure of MOFs allows for property tuning through variations in metal ions and organic ligands, making them promising candidates for energy and environmental applications. Among various MOFs, Zn-MOF stands out for its ease of synthesis, stable zinc coordination properties, and potential for high-temperature applications [4]. However, the reality shows that many Zn-MOFs still have limitations in thermal stability, and the factors controlling this stability have not been fully quantified. This gap hampers the optimal utilization of Zn-MOFs in catalytic and clean energy applications, which require resistance to extreme conditions.

To address this challenge, this study focuses on the development of a machine learning-based predictive model in the form of stacking to predict the thermal stability of Zn-MOFs. By integrating multiple base regressors into a single meta-learner, stacking allows for robust and accurate predictions, even in noisy or heterogeneous datasets. Based on the Quantitative Structure-Property Relationship (QSPR) paradigm, this model utilizes structural descriptors to capture the intrinsic relationship between molecular architecture and thermal behavior [5] [6]. According to the reviewed literature, no prior research has applied stacking ensemble methods for predicting the thermal stability of Zn-MOFs, which constitutes the principal novelty of this study. Through this approach, the study aims to generate accurate predictions that reduce reliance on time-consuming experimental methods. Consequently, the findings are expected to accelerate the design of thermally stable MOF materials, support the implementation of sustainable energy technologies, and contribute to the global climate change mitigation agenda.

The application of machine learning in materials science has introduced new approaches to accelerate the discovery of functional materials [7]. This is particularly relevant for systems where experimental evaluations demand

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1390−1399
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8329

significant resources. These materials are known for their distinctive properties, such as high porosity, considerable surface area, and chemical design flexibility, making them prime candidates for applications like gas separation, membrane development, and corrosion protection. However, the complexity of crystalline architectures and the vast number of potential chemical combinations present significant challenges, necessitating predictive, high-performance computational approaches to support and guide experiments more efficiently [8], [7]. Within this framework, research that integrates machine learning with the study of MOFs carries significant scientific value, as it not only deepens fundamental understanding in materials science but also provides practical relevance to the energy field by supporting the development of more stable and environmentally adaptive energy storage systems.

A key study in the application of machine learning (ML) to materials science was conducted by Zhang et al. (2022) [9]. This study built a large-scale database containing 8,167 ionic liquid@MOF (IL@MOF) composites. Using the Random Forest (RF) algorithm, the study successfully predicted membrane performance for CO_2 capture. Further analysis showed that accessible volume (AV) and gravimetric surface area (gASA) were key structural descriptors determining membrane selectivity and permeability. These findings highlight how ML can be used to efficiently screen material candidates before the synthesis stage, saving time and experimental costs.

A similar study was conducted by Usman et al. (2024), who applied Gaussian Process Regression (GPR) to evaluate the separation efficiency of UiO-66-NH_2-based membranes in oil wastewater treatment [10]. Their results showed that GPR outperformed Support Vector Machine and Decision Tree models, especially when combined with feature engineering to highlight the most relevant descriptors. This emphasizes that the integration of ML with experimental design can strengthen the analysis in complex separation systems.

Furthermore, Daglar and Keskin (2022) combined ML approaches with molecular simulations to predict gas adsorption and diffusion properties on thousands of MOFs and mixed MOF/polymer membranes (MMM) [11]. By testing over 31,000 systems using various ML models, they gained comprehensive insights into the potential of MOF membranes for gas separation applications. This study demonstrated the scalability of ML in handling large material datasets and underscored its role in evaluating hypothetical MOFs that have yet to be synthesized.

Another example is provided by Liang et al. (2021), who used the Extreme Gradient Boosting (XGBoost) algorithm to identify key structural descriptors in the separation of Xe/Kr gases [12]. The analysis showed that simple geometric features such as density, porosity, and pore volume were sufficient to produce predictions with high accuracy. XGBoost not only surpassed traditional models in prediction and generalization but also achieved an R^2 value greater than 0.95, underscoring the effectiveness of ML-based feature approaches.

Additionally, Anadebe et al. (2022) explored the Adaptive Neuro-Fuzzy Inference System (ANFIS) to predict the corrosion inhibition performance of cerium-based MOF [13]. This hybrid model was able to capture complex non-linear interactions between experimental input variables and performance output, yielding an almost perfect correlation with empirical data. This demonstrates the high potential of ANFIS as an AI approach capable of reliably representing complex electrochemical phenomena.

Previous studies have shown that the application of machine learning (ML), primarily through the integration of feature engineering and simulation data, can significantly enhance predictive capabilities in analyzing Metal–Organic Frameworks (MOF)-based systems. However, there exists a significant research gap in the literature, particularly concerning the application of ensemble stacking methods to predict the thermal stability of MOFs. Most studies still focus on single models, which are often suboptimal in handling outliers and data distribution variations in experimental datasets. Moreover, although the applications of MOFs in gas separation, membrane performance, and corrosion inhibition have been extensively reported, the aspect of thermal stability crucial for the utilization of MOFs under high-temperature conditions such as in catalysis and energy storage has been relatively underexplored, particularly in the context of predictive modeling based on Zn-MOFs.

To bridge this gap, this study proposes the development of a Quantitative Structure Property Relationship (QSPR) model based on stacking regression to predict the thermal stability of Zn-MOFs. This model is constructed by leveraging structural descriptors as features representing the chemical characteristics and atomic interactions of Zn-MOF compounds. Through the integration of multiple basic regression algorithms into a unified learning pipeline, the stacking approach is expected to enhance predictive robustness, minimize the influence of data variation and outliers, and contribute significantly to the accelerated identification of thermally stable Zn-MOF materials for industrial applications [13].

2. RESEARCH METHODOLOGY

This study applies a methodological framework to analyze the thermal characteristics of Zn-MOF compounds and develop an accurate predictive model based on the structural components that influence their stability. As shown in Figure 1, the framework begins with a dataset of 151 Zn-MOF compounds, each described by four key structural descriptors. The dataset is processed through data pre-processing and exploratory analysis, then divided into training and testing subsets. The modeling stage employs a stacking ensemble regression approach, integrating multiple regression algorithms to enhance predictive accuracy and robustness. The final outcome is a predictive model evaluated on the testing data, providing a comprehensive assessment of its accuracy and reliability in predicting the thermal stability of Zn-MOF compounds.

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1390−1399
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8329

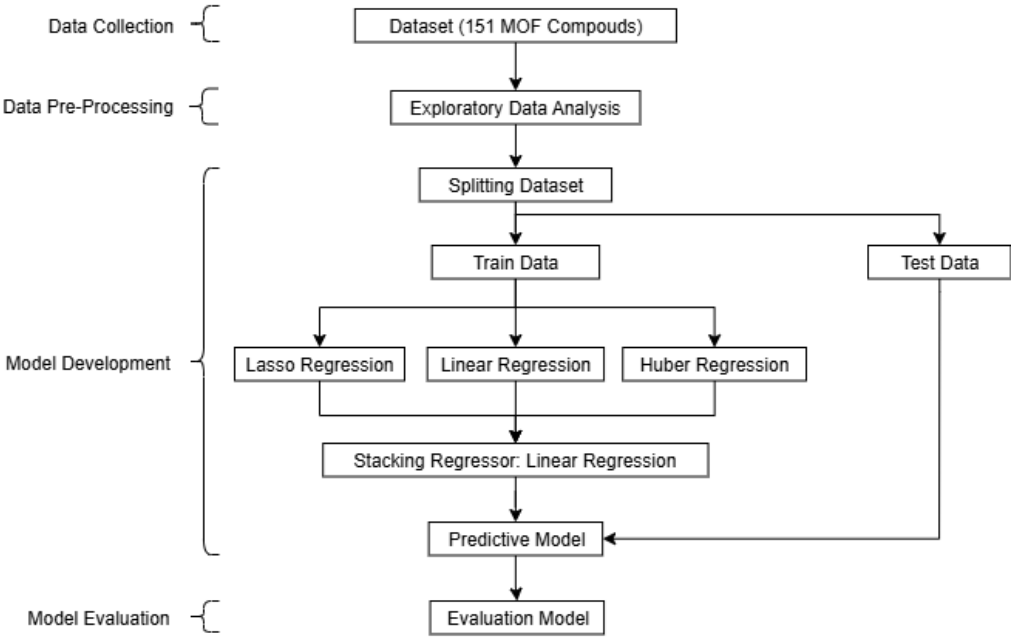

Figure 1. Research Framework

2.1 Pre-Processing Data

The first step in developing a machine learning model to predict the thermal stability of Zn-MOFs is Data Pre-processing, which involves steps such as data cleaning to detect and address missing values, and identifying and managing outliers to prevent impact on the prediction results. Given the structural diversity present in Zn-MOF compounds, handling outliers becomes crucial to avoid model distortion due to extreme values [16].

Additionally, data normalization was performed using the Robust Scaler to ensure balanced feature scales across the descriptors. Alternative feature scaling techniques such as the Min-Max Scaler and Standard Scaler were also considered; however, both methods are highly sensitive to outliers and may reduce the reliability of the predictive model. By contrast, the Robust Scaler employs the median and interquartile range (IQR), making it more resistant to extreme values and better suited for heterogeneous data distributions. This choice ensures that each feature contributes proportionally to the model without being dominated by anomalies. The overall goal of this pre-processing step is to produce a clean, consistent, and well-scaled dataset, thereby enabling the development of an accurate and reliable predictive model [16]

2.2 Exploratory Data Analysis (EDA)

Following the pre-processing stage, Exploratory Data Analysis (EDA) was performed to gain a deeper understanding of the dataset's characteristics and to evaluate the effectiveness of the applied transformations. Boxplots were employed to illustrate the distribution of each descriptor, confirming the presence of extreme values while also demonstrating that the scaling procedure reduced their disproportionate influence. Histograms provided additional insights into the distributional properties of the descriptors, revealing variations ranging from approximately symmetric patterns to skewed distributions. This further reinforced the importance of normalization to ensure that each feature contributes proportionally to the modeling process. In addition, the correlation heatmap depicted the degree of association among descriptors, with several variable pairs exhibiting moderate to high correlations. Such information is essential for anticipating potential multicollinearity, which could affect the stability and interpretability of regression models. Through these visual analyses, EDA not only complemented the results of pre-processing but also provided an empirical basis for feature selection and model development. By confirming data quality, distributional balance, and inter-feature relationships, EDA ensured that the modeling process was built upon a dataset that is both reliable and representative [17].

2.3 Model Development

This study develops a predictive model to forecast the thermal stability of Zn-MOFs using a stacking approach. The three base algorithms used in this model are Linear Regression, Lasso Regression, and Huber Regression [18], [19]. The selection of these algorithms is based on their ability to handle data with high variation and complexity, as well as their capacity to learn the relationships between structural descriptors and the target thermal stability (TS). Each base model is optimized through hyperparameter tuning to ensure optimal performance on the training data.

Once the base models are trained, the prediction results from each model are combined using a stacking approach, with Linear Regression serving as the meta-model. This meta-model integrates the predictions from the

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1390−1399
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8329

base models to generate a more accurate and reliable output. The stacking approach aims to enhance the model's generalization ability by leveraging the strengths of each base model to capture complex patterns in the Zn-MOF data.

2.4.1 Linear Regression

Linear Regression is used to build the base model that captures the linear relationship between the descriptor variables and the thermal stability of Zn-MOFs. This model works by minimizing prediction errors, providing a robust yet straightforward foundation for predicting thermal stability. In the stacking approach, the prediction results from Linear Regression are combined with those from other models to enhance the overall prediction accuracy [19].

2.4.2 Lasso Regression

Lasso Regression is a variant of Linear Regression that adds L1 regularization to shrink the coefficients of less relevant variables. This approach helps select the descriptors most influential to thermal stability, thereby reducing the risk of overfitting and improving the model's generalization. In the stacking approach, Lasso Regression serves to enhance the predictions by selecting the optimal features [19].

2.4.3 Huber Regression

Huber Regression is a method that is more robust to outliers, combining the advantages of Mean Squared Error (MSE) and Mean Absolute Error (MAE). This model is effective in handling data containing extreme values, such as in the Zn-MOF dataset, by providing stable predictions even in the presence of anomalous data. In the stacking approach, Huber Regression strengthens the overall stability of the model by effectively managing outliers [18].

2.4.4 Stacking Regression

This study applies the stacking ensemble method by integrating the prediction results from several base models to enhance overall prediction performance, as shown in Figure 2.

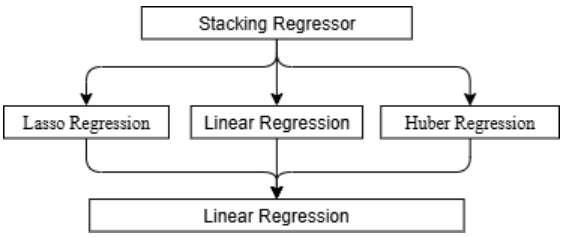

Figure 2. Stacking-Based Regressor Framework

2.4 Model Evaluation

In evaluating prediction results, three key metrics are commonly used to assess the performance of the model: RMSE (Root Mean Squared Error), R2 (R-squared), and MAE (Mean Absolute Error). RMSE, as presented in Equation (1), calculates the square root of the average of the squared differences between the actual values (y_i) and the predicted values ($\hat{y}_i$). This metric quantifies the magnitude of prediction error, where n represents the total number of data points.

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

R2, shown in Equation (2), measures the proportion of variance in the actual values that can be explained by the model. It is computed by dividing the sum of squared differences between the actual values (y_i) and their predicted values ($\hat{y}_i$) by the sum of squared differences between the actual values (y_i) and the mean of the actual values ($\bar{y}$).

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

Lastly, MAE, as defined in Equation (3), computes the average of the absolute differences between the actual values (y_i) and the predicted values ($\hat{y}_i$). This metric provides insight into the overall accuracy of the predictions [20].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (3)$$

3. RESULT AND DISCUSSION

This section presents an analysis of the predicted thermal stability of Zn-MOFs to evaluate the accuracy and performance of the model in predicting the thermal stability of Metal-Organic Frameworks (MOFs). This study focuses on the results of the Exploratory Data Analysis (EDA) and the evaluation of the predictive model that employs the stacking technique.

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1390−1399
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8329

3.1 Results of Exploratory Data Analysis (EDA)

In this subsection, an exploratory data analysis is conducted to understand the characteristics of the dataset used, including descriptive statistical analysis, identification of outliers, skewness, and correlation between descriptors and thermal stability (TS).

3.1.1 Descriptive Statistics

Descriptive statistical analysis aims to understand the distribution and variability of the data for each descriptor used in the modeling, as well as the target variable, which is thermal stability (TS).

Table 1. Descriptive Statistics

Descriptor	Mean	Variance	Standard Deviation
nN	6.087616	57.820970	7.604010
nZn	3.353179	65.619692	8.100598
Het	-4.149854	106.838406	10.336267
Lig	2.487258	1.926191	1.387873
TS	2.574437	0.008166	0.090367

Based on the calculations presented in Table 1, the number of nitrogen atoms (nN) exhibits an average value of 6.08, with a variance of 57.82 and a standard deviation of 7.60. These results indicate that the presence of nitrogen atoms within the MOF structure varies considerably across samples. This variation reflects differences in the types and quantities of N-donor ligands employed, leading to a relatively broad data distribution for this parameter.

Meanwhile, the number of zinc atoms (nZn) has an average value of 3.35, with a variance of 65.61 and a standard deviation of 8.10. Although the average zinc content is smaller compared to nN, the data spread is also high. This suggests a diversity in the number of metal centers in the MOF structure, contributing to variations in coordination and material topology.

The descriptor for heteroatom interactions (Het) shows the most fluctuating characteristics, with an average of -4.15, a variance of 106.83, and the highest standard deviation of 10.34. The negative average value indicates that, in most MOF structures, heteroatom interactions tend to be destabilizing. However, the vast data spread also suggests a complex role for Het, potentially either enhancing or reducing thermal stability, making this variable likely to have a significant non-linear effect in the modeling process.

In contrast to Het, the ligand descriptor (Lig) exhibits a relatively homogeneous distribution, with an average of 2.49, a variance of 1.93, and a standard deviation of 1.39. This indicates that the numerical representation of ligands in the dataset does not exhibit significant variation, suggesting a more consistent role of ligands in thermal stability compared to other descriptors. The target variable, thermal stability (TS), has an average value of 2.57, with a minimal variance of 0.008 and a standard deviation of 0.09. This indicates that TS values are relatively uniform across all samples, despite the high variation in the input descriptors.

3.1.2 Outlier Identification

Outlier identification is a crucial step in data analysis to detect values that significantly differ from the majority of the other data. The presence of outliers can impact statistical analysis results and model predictions.

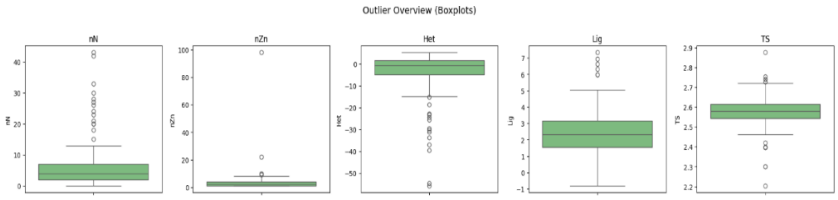

Figure 3. Outlier Visualization of All Variables using Boxplot

As illustrated in Figure 3, outlier identification through the boxplot reveals that the nN (number of nitrogen atoms) variable contains a substantial number of extreme values above the main distribution. This indicates the presence of MOFs with nitrogen content considerably higher than the mean. For nZn (number of zinc atoms), several outliers were also detected, including extremely high values, suggesting the occurrence of large metal clusters or complex frameworks. The Het (heteroatom interactions) descriptor exhibits numerous negative outliers, reflecting weak interactions or potential destabilization in certain MOFs. This finding is particularly important, as such anomalies may exert a non-linear influence on thermal stability. In contrast, the Lig (ligand fragments) descriptor shows relatively few high-value outliers, likely associated with ligands of unique structures, while maintaining a more homogeneous distribution compared to nN and Het.

The TS (thermal stability) target appears relatively uniform, but there are still some outliers on both the lower and upper ends of the distribution. These extreme values indicate the presence of MOFs with thermal stability lower or higher than that of the majority of the samples, which may affect the performance of the predictive model. Overall,

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1390−1399
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8329

the Zn-MOF dataset contains several extreme values, particularly in nN, nZn, and Het. These outliers are not merely considered as data errors but also reflect the diversity of MOF structures and intrinsic properties.

3.1.3 Skewness Identification

Skewness analysis is performed to measure the extent to which the data distribution deviates from a normal distribution, either to the left (negative skew) or to the right (positive skew). Assessing skewness is essential as it can influence the choice of modeling techniques and the interpretation of results (Figure 4).

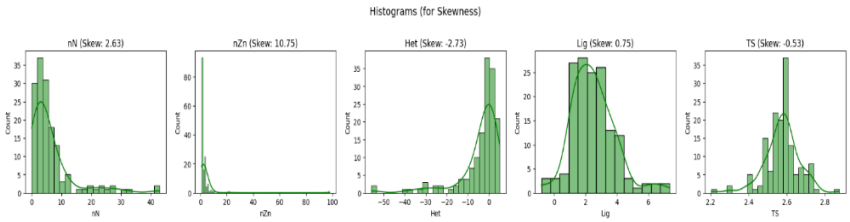

Figure 4. Skewness Visualization of All Variables using Histogram

Skewness analysis is performed to assess the skewness of the data distribution for each descriptor and the target (TS). The measurement results show that the nN (number of nitrogen atoms) variable has a skewness value of 2.63, indicating that its distribution is heavily skewed to the right (positive skew). This condition suggests that most of the nN data is concentrated at lower values, while there are a few MOFs with very high nitrogen content. For the nZn (number of zinc atoms) variable, the skewness is even more extreme, with a value of 10.75. This distribution is strongly skewed to the right, indicating that the majority of samples have a low number of Zn atoms, but a few outliers with much higher Zn content cause the distribution to be highly asymmetric.

In contrast to nN and nZn, the Het (heteroatom interactions) descriptor exhibits negative skewness with a value of -2.73, indicating a strong skew to the left. This means that most MOFs have heteroatom interaction values close to zero or positive, while a small number of samples show extreme negative interaction values, indicating a tendency toward destabilization. The Lig (ligand fragments) descriptor is relatively more balanced, with a skewness value of 0.75, suggesting a slight rightward skew but still close to normal. This indicates that the numerical representation of ligands is relatively homogeneous. Meanwhile, the TS (thermal stability) target has a skewness value of -0.53, indicating that its distribution is relatively symmetrical with a slight leftward skew. This nearly normal distribution suggests that the thermal stability of the MOF in the dataset is not heavily biased, although. However, there are a few samples with TS values lower than the majority.

3.1.4 Correlation Between Descriptors and Thermal Stability

Correlation analysis aims to evaluate the extent of the linear relationships among the descriptors used in the modeling, as well as the relationship between the descriptors and thermal stability (TS). The correlation between these variables can provide insights into the strength of each descriptor's influence on thermal stability, as well as how the descriptors interact with each other. Identifying significant correlations is essential for understanding the dynamics between variables in the Metal-Organic Framework (MOF) structure and for determining which descriptors contribute the most in predicting thermal stability.

Correlation analysis in Figure 5 was performed to evaluate the linear relationships between each descriptor and thermal stability (TS), as well as the relationships among the descriptors themselves. Based on the correlation heatmap, it is observed that the nN (number of nitrogen atoms) variable has a robust and negative correlation with Het (heteroatom interactions), with a value of (-0.98). This nearly perfect relationship indicates a close connection between the number of nitrogen atoms in the structure and the characteristics of heteroatom interactions, where an increase in nitrogen content tends to be followed by a decrease in Het values.

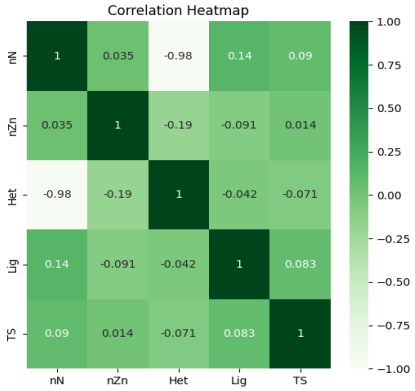

Figure 5. Correlation Between Variables and Thermal Stability

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1390−1399
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8329

Meanwhile, correlations among other descriptors are generally relatively weak. nZn (number of zinc atoms) only shows a low correlation with nN (0.035) and Het (-0.19), indicating that the variation in zinc content is not entirely dependent on the number of nitrogen atoms or the strength of heteroatom interactions. The Lig (ligand fragments) descriptor also has low correlations with other variables, including TS (0.083), which means that the numerical variation in ligands within the dataset is relatively independent of other descriptors.

When considering TS (thermal stability) as the target, all descriptors show relatively low linear correlations. The highest correlation values are observed for nN (0.09) and Lig (0.083), while Het shows a slight negative correlation (-0.071) and nZn is almost unrelated (0.014). This suggests that no single descriptor has a strong linear influence on thermal stability. Therefore, the thermal stability of MOFs is likely influenced by a combination of complex interactions among variables, rather than by a single dominant factor.

Overall, the results of the EDA indicate that the Zn-MOF dataset exhibits a high degree of heterogeneity, as reflected in the large variances, the presence of outliers, and the asymmetric distributions observed in several descriptors. These characteristics can negatively impact the performance of conventional regression models, particularly due to their sensitivity to extreme values and their limited ability to accommodate imbalanced data distributions. Furthermore, the weak linear correlations between descriptors and thermal stability suggest that no single variable exerts a dominant influence, with predictions instead being shaped by the combined interactions among multiple factors. Accordingly, the EDA findings justify the adoption of stacking ensemble regression, as this approach demonstrates greater adaptability to data heterogeneity and is capable of producing more accurate, reliable, and robust predictions of Zn-MOF thermal stability.

3.2 Model Development

This study develops a predictive model for analyzing the thermal stability of Zn-MOF using the stacking method. Linear Regression (LR), Lasso Regression, and Huber Regression are selected as base models, each offering a different approach to the relationship between the descriptors (nN, nZn, Het, Lig) and thermal stability (TS). Stacking Regression is then used to combine the predictions from these three base models, to improve prediction accuracy and generalization.

3.2.1 Linear Regression

The Linear Regression (LR) model is used as a base model to evaluate the extent to which the linear relationship between the descriptors (nN, nZn, Het, and Lig) can explain the variation in the target thermal stability (Table 2).

Table 2. Comparison of Evaluation Metrics in the Linear Regression Model

Dataset	MAE	RMSE	R^2
Training Set	*0.001904	*0.002277	*0.999406
Testing Set	0.002937	0.005452	0.994761

Note: *) Best Performance

On the training data, the Linear Regression model yields very low prediction errors, with an MAE of 0.001904 and an RMSE of 0.002277. The R^2 value of 0.999406 indicates that the model can explain more than 99% of the variation in the training data. On the test data, the model's performance slightly decreases, with an MAE of 0.002937 and an RMSE of 0.005452. The R^2 value also decreases to 0.994761. Nevertheless, this value still indicates a very high level of accuracy. The slight difference between the training and test data suggests that the Linear Regression model does not suffer from overfitting and demonstrates good generalization ability.

3.2.2 Lasso Regression

The Lasso Regression model is used as an alternative to Linear Regression by adding L1 regularization, which serves to shrink the coefficients of less influential variables, thereby reducing the risk of overfitting and improving the model's generalization on the test data (Table 3).

Table 3. Comparison of Evaluation Metrics in the Lasso Regression Model

Dataset	MAE	RMSE	R^2
Training Set	*0.001898	*0.002279	*0.999405
Testing Set	0.002918	0.005390	0.994878

Note: *) Best Performance

On the training data, the performance of Lasso Regression is very high, with an MAE of 0.001898 and an RMSE of 0.002279. The R^2 value of 0.999405 indicates that the model can explain almost the entire variation in the training data. These results are nearly identical to those of the Linear Regression model. On the test data, Lasso provides stable results with an MAE of 0.002918, an RMSE of 0.005390, and an R^2 of 0.994878. These values are slightly better than those of Linear Regression, indicating that regularization helps improve the model's generalization on new data.

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1390−1399
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8329

3.2.3 Huber Regression

The Huber Regression model is used because of its robustness to outliers. Huber Regression combines a squared loss function for minor errors and an absolute loss function for significant errors, making it more stable in the presence of extreme data (Table 4).

Table 4. Comparison of Evaluation Metrics in the Huber Regression Model

Dataset	MAE	RMSE	R^2
Training Set	*0.001869	*0.002294	*0.999397
Testing Set	0.002914	0.005473	0.994720

Note: *) Best Performance

On the training data, Huber Regression performs very well with an MAE of 0.001869, an RMSE of 0.002294, and an R^2 of 0.999397. This indicates that the model can learn the relationships between variables with high accuracy. On the test data, the model maintains good performance with an R^2 of 0.994720. Although the RMSE is slightly higher compared to Lasso, the advantage of Huber lies in its ability to handle outliers, making it more stable when dealing with data containing extreme values.

3.2.4 Stacking Regression

This approach leverages the strengths of each base model through a meta-learner that optimizes the combination of predictions, allowing for a more comprehensive capture of the complex relationships among variables [21] (Table 5).

Table 5. Comparison of Evaluation Metrics in the Stacking Model

Dataset	MAE	RMSE	R^2
Training Set	*0.001914	*0.002281	*0.999404
Testing Set	0.002915	0.005262	0.995119

Note: *) Best Performance

On the training set, the stacking model achieves performance with an MAE of 0.001914, an RMSE of 0.002281, and an R^2 of 0.999404. These values are comparable to both the linear and robust models that serve as its base, with minimal prediction errors and an explanation of over 99.9% of the data variation. On the testing set, the model maintains excellent performance with an MAE of 0.002915, an RMSE of 0.005262, and an R^2 of 0.995119. The higher R^2 value compared to the base models indicates that stacking improves generalization. Additionally, the slight difference between the training and testing results suggests that the model does not exhibit signs of overfitting.

3.3 Model Evaluation

This subsection provides a detailed comparison of the evaluation results of the regression models that have been developed, to determine which model is the most optimal for predicting thermal stability and is capable of handling the variability and complexity of the data (Table 6).

Table 6. Comparison of Evaluation Metrics Across All Models

	Training Set			Testing Set		
	MAE	RMSE	R2	MAE	RMSE	R^2
Linear Regression	0.001904	*0.002277	*0.999406	0.002937	0.005452	0.994761
Lasso Regression	0.001898	0.002279	0.999405	0.002918	0.005390	0.994878
Huber Regression	*0.001869	0.002294	0.999397	*0.002914	0.005473	0.994720
Stacking Regression	0.001914	0.002281	0.999404	0.002915	*0.005262	*0.995119

Note: *) Best Performance

Linear Regression delivers strong performance with an R^2 value of 0.994761 and an RMSE of 0.005452 on the test data. This indicates that the linear relationship can explain most of the target variation. However, its dependence on linearity makes it less flexible when dealing with more complex patterns. Lasso Regression offers a slight improvement over Linear Regression. With an R^2 value of 0.994878 and an RMSE of 0.005390 on the test data, Lasso is more stable than Linear Regression. This shows that the L1 regularization plays a role in enhancing generalization by reducing the influence of less relevant variables without sacrificing main accuracy.

Huber Regression provides a different dynamic. On the training data, Huber achieves the lowest MAE of 0.001869, indicating high accuracy in learning the data patterns. However, on the test data, its performance is slightly lower compared to Linear and Lasso Regression, with an R^2 of 0.994720 and an RMSE of 0.005473. This performance indicates a trade-off: pure accuracy is somewhat reduced, but the model becomes more robust to extreme values. Given that the Zn-MOF dataset contains several outliers, this advantage remains relevant, even though Huber's evaluation metrics are not the best.

On the other hand, Stacking Regression clearly demonstrates the most outstanding performance. With an R^2 value of 0.995119 and an RMSE of 0.005262 on the test data, stacking not only matches the base models' performance

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1390−1399
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8329

but also surpasses them. A comparison of the numbers shows that the RMSE of stacking is lower than that of Linear Regression (0.005262 vs 0.005452), lower than Lasso Regression (0.005262 vs 0.005390), and also lower than Huber Regression (0.005262 vs 0.005473). Additionally, the R^2 value of stacking is the highest among all models. This indicates that the ensemble approach successfully combines the strengths of Linear Regression in capturing basic relationships, Lasso in providing stability, and Huber in handling outliers.

In other words, this comparison shows that Linear Regression performs adequately, Lasso Regression is slightly better in terms of stability, and Huber Regression is relevant for data with extreme values. However, Stacking Regression integrates the advantages of all three and produces the most optimal generalization model. The stacking model's position as the best-performing model is clearly reflected in the slight yet consistent differences across all metrics, making it the most suitable model for predicting the thermal stability of Zn-MOFs.

3.4 External Validation

To test the reliability and external validity of the developed model, a comparison was made between the experimental results and the predictions obtained using the Stacking Regression method. Validation against experimental data is an important step, as it provides additional evidence regarding the consistency and reliability of the model. The data used comes from the Zn-MOF study by Wang et al. (2019), with the primary descriptors being nN, nZn, Het, and Lig [22].

Table 7. Comparison of Experimental TS Prediction Results with Stacking

Descriptor				Log TS	
nN	nZn	Het	Lig	Experimental [22]	Prediction
2	3	2,66	4,13	2,720	2,716

The results show (Table 7) that the experimental value of logTS, 2.720, is very close to the predicted value from the stacking model, 2.716. The difference between the two is only 0.004, indicating a very high level of agreement. This closeness underscores that the stacking model not only provides high accuracy but also demonstrates the ability to represent experimental results with good consistency. This finding has important implications in materials research, as the stacking-based machine learning predictions can closely match experimental results with very high precision. Thus, this method can be seen as a much more efficient alternative in terms of time and cost compared to traditional laboratory experimental methods.

4. CONCLUSION

This study demonstrates that the machine learning-based stacking ensemble approach is a highly effective strategy for predicting the thermal stability of Zn-MOFs. By combining three base models, namely Linear Regression, Lasso Regression, and Huber Regression into a Linear Regression meta-model, optimal predictive performance is achieved compared to a single model. Evaluation results on the test data, with an MAE of 0.002915, RMSE of 0.005262, and R^2 of 0.995119, show extremely high prediction accuracy and reliable generalization ability. These findings confirm that stacking effectively integrates the strengths of each base algorithm, resulting in more stable predictions that are resilient to variability and outliers in the data. In addition to achieving superior accuracy, this method also offers significant time and cost efficiency compared to conventional experimental procedures. By utilizing structural descriptors such as the number of zinc atoms, the number of nitrogen atoms, heteroatom interactions, and ligand fragments, this stacking-based QSPR model successfully represents the intrinsic relationship between chemical structure and thermal stability. This opens up significant opportunities to accelerate the identification and design process of MOFs capable of withstanding high-temperature conditions, which is crucial for industrial energy and catalysis applications. However, the limitation of this study lies in the relatively small dataset size, so the potential for generalization to a broader range of MOF structures has not yet been thoroughly tested. Therefore, further validation using a larger and more diverse dataset is required to strengthen the model's reliability. For future development, the study could explore the application of other ensemble techniques, such as bagging or blending, as well as integration with non-linear models, to test the potential for further improvement in prediction accuracy. Additionally, enriching the dataset through new experimental results or simulation data will broaden the model's scope and enhance its scalability. Thus, this study underscores that the stacking technique not only provides highly accurate predictions but also serves as an efficient and practical approach to support the discovery and optimization of thermally stable MOF materials, in line with the needs of clean and sustainable energy technologies.

ACKNOWLEDGMENT

We sincerely thank the Directorate of Learning and Student Affairs, Directorate General of Higher Education, Ministry of Education, Science, and Technology of the Republic of Indonesia, for the funding provided under the PKM Grant No. 1995/B2/DT.01.00/2025, which has been crucial for the successful completion of this research. We also acknowledge Universitas Dian Nuswantoro (UDINUS) and the Research Center for Quantum Computing and

Building of Informatics, Technology and Science (BITS)
Volume 7, No 2, September 2025 Page: 1390−1399
ISSN 2684-8910 (media cetak)
ISSN 2685-3310 (media online)
DOI 10.47065/bits.v7i2.8329

Materials Informatics, Study Program in Informatics Engineering, Faculty of Computer Science, Universitas Dian Nuswantoro, Indonesia, for their support, facilities, and valuable guidance throughout this study.

REFERENCES

[1] "CO2 Emissions – Global Energy Review 2025 – Analysis," IEA. Accessed: Sept. 22, 2025. [Online]. Available: https://www.iea.org/reports/global-energy-review-2025/co2-emissions

[2] N. Liang and Y. Zhao, "A review on thermal stability of nanostructured materials," *J. Alloys Compd.*, vol. 938, p. 168528, 2023, doi: https://doi.org/10.1016/j.jallcom.2022.168528.

[3] G. Cai, P. Yan, L. Zhang, H.-C. Zhou, and H.-L. Jiang, "Metal–Organic Framework-Based Hierarchically Porous Materials: Synthesis and Applications," *Chem. Rev.*, vol. 121, no. 20, pp. 12278–12326, 2021, doi: 10.1021/acs.chemrev.1c00243.

[4] G. Valdebenito, M. Gonzaléz-Carvajal, L. Santibañez, and P. Cancino, "Metal–Organic Frameworks (MOFs) and Materials Derived from MOFs as Catalysts for the Development of Green Processes," *Catalysts*, vol. 12, no. 2, 2022, doi: 10.3390/catal12020136.

[5] M. Yu *et al.*, "Quantitative structure-property relationship (QSPR) framework assists in rapid mining of highly Thermostable polyimides," *Chem. Eng. J.*, vol. 465, p. 142768, 2023, doi: https://doi.org/10.1016/j.cej.2023.142768.

[6] S. Datta, V. A. Dev, and M. R. Eden, "Developing Non-linear Rate Constant QSPR using Decision Trees and Multi-Gene Genetic Programming," in *13th International Symposium on Process Systems Engineering (PSE 2018)*, vol. 44, M. R. Eden, M. G. Ierapetritou, and G. P. Towler, Eds., in Computer Aided Chemical Engineering, vol. 44. , Elsevier, 2018, pp. 2473–2478. doi: https://doi.org/10.1016/B978-0-444-64241-7.50407-9.

[7] M. Akrom, S. Rustad, A. G. Saputro, A. Ramelan, F. Fathurrahman, and H. K. Dipojono, "A combination of machine learning model and density functional theory method to predict corrosion inhibition performance of new diazine derivative compounds," *Mater. Today Commun.*, vol. 35, p. 106402, June 2023, doi: 10.1016/j.mtcomm.2023.106402.

[8] M. M. Hossain and K. Roy, "QSPR modeling of thermal stability of reactive and self-reactive chemicals using 2D descriptors: Predictions of heat of decomposition, Self-Accelerating decomposition temperature, and onset temperature," *Mater. Today Commun.*, vol. 46, p. 112751, 2025, doi: https://doi.org/10.1016/j.mtcomm.2025.112751.

[9] Z. Zhang *et al.*, "Machine learning aided high-throughput prediction of ionic liquid@MOF composites for membrane-based CO2 capture," *J. Membr. Sci.*, vol. 650, p. 120399, 2022, doi: https://doi.org/10.1016/j.memsci.2022.120399.

[10] J. Usman, S. I. Abba, N. Baig, N. Abu-Zahra, S. W. Hasan, and I. H. Aljundi, "Design and Machine Learning Prediction of In Situ Grown PDA-Stabilized MOF (UiO-66-NH_2) Membrane for Low-Pressure Separation of Emulsified Oily Wastewater," *ACS Appl. Mater. Interfaces*, vol. 16, no. 13, pp. 16271–16289, Apr. 2024, doi: 10.1021/acsami.4c00752.

[11] H. Daglar, S. Aydin, and S. Keskin, "MOF-based MMMs Breaking the Upper Bounds of Polymers for a Large Variety of Gas Separations," *Sep Purif Technol*, vol. 281, p. 119811, 2022.

[12] H. Liang, K. Jiang, T.-A. Yan, and G.-H. Chen, "XGBoost: An Optimal Machine Learning Model with Just Structural Features to Discover MOF Adsorbents of Xe/Kr," *ACS Omega*, vol. 6, no. 13, pp. 9066–9076, 2021, doi: 10.1021/acsomega.1c00100.

[13] V. C. Anadebe, V. I. Chukwuike, S. Ramanathan, and R. C. Barik, "Cerium-based metal organic framework (Ce-MOF) as corrosion inhibitor for API 5L X65 steel in CO2- saturated brine solution: XPS, DFT/MD-simulation, and machine learning model prediction," *Process Saf. Environ. Prot.*, vol. 168, pp. 499–512, 2022, doi: https://doi.org/10.1016/j.psep.2022.10.016.

[14] H. A. Azies, M. Akrom, S. Rustad, and H. K. Dipojono, "Robust Machine Learning for Predicting Thermal Stability of Metal-Organic Framework," *Chem. Afr.*, vol. 7, no. 8, pp. 4669–4681, Oct. 2024, doi: 10.1007/s42250-024-01080-4.

[15] M. Moharramnejad, L. Tayebi, A. R. Akbarzadeh, and A. Maleki, "A simple, robust, and efficient structural model to predict thermal stability of zinc metal-organic frameworks (Zn-MOFs): The QSPR approach," *Microporous Mesoporous Mater.*, vol. 336, p. 111815, May 2022, doi: 10.1016/j.micromeso.2022.111815.

[16] C. Ghosh, "Data Pre-processing," in *Data Analysis with Machine Learning for Psychologists: Crash Course to Learn Python 3 and Machine Learning in 10 hours*, Cham: Springer International Publishing, 2022, pp. 55–85. doi: 10.1007/978-3-031-14634-3_3.

[17] K. Purohit, "Separation of Data Cleansing Concept from EDA," *Int. J. Data Sci. Anal.*, vol. 7, no. 3, p. 89, 2021, doi: 10.11648/j.ijdsa.20210703.16.

[18] Y. Feng and Q. Wu, "A statistical learning assessment of Huber regression," *J. Approx. Theory*, vol. 273, p. 105660, Jan. 2022, doi: 10.1016/j.jat.2021.105660.

[19] J. O. Ogutu, T. Schulz-Streeck, and H.-P. Piepho, "Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions," *BMC Proc.*, vol. 6, no. 2, p. S10, May 2012, doi: 10.1186/1753-6561-6-S2-S10.

[20] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput. Sci.*, vol. 7, p. e623, July 2021, doi: 10.7717/peerj-cs.623.

[21] T. A. Munshi, K. Popi, L. N. Jahan, M. F. Howladar, and M. Hashan, "Stacking modeling with genetic algorithm-based hyperparameter tuning for uniaxial compressive strength prediction," *Appl. Comput. Geosci.*, vol. 27, p. 100276, Sept. 2025, doi: 10.1016/j.acags.2025.100276.

[22] X.-Q. Wang *et al.*, "A water-stable zinc(II)–organic framework as a multiresponsive luminescent sensor for toxic heavy metal cations, oxyanions and organochlorine pesticides in aqueous solution," *Dalton Trans.*, vol. 48, no. 44, pp. 16776–16785, 2019, doi: 10.1039/C9DT03195B.