BAREKENG: Journal of Mathematics and Its Applications
June 2026 Volume 20 Issue 2 Page 1373–1388
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss2pp1373-1388

EVALUATING THE EFFECTIVENESS OF KERNEL EXTREME LEARNING MACHINES OVER CONVENTIONAL ELM FOR AIR QUALITY INDEX PREDICTION

Meta Kallista [1], Ig. Prasetya Dwi Wibawa [2*], Sultan Chisson Obie [3]

[1,3]*Computer Engineering Study Program, School of Electrical Engineering, Telkom University*
[2]*Electrical Engineering Study Program, School of Electrical Engineering, Telkom University*
[1]*Center of Excellence for Human Centric Engineering, Institute of Sustainable Society, Telkom University*
[2]*Center of Excellence for Smart Transportation and Robotics,*
Research Institute for Connectivity and Convergence for Smart Living, Telkom University
Jln. Telekomunikasi no. 1, Bandung, Jawa Barat, 40257, Indonesia

*Corresponding author's e-mail: * **prasdwibawa@telkomuniversity.ac.id***

Article Info

Article History:

Received: 2nd June 2025
Revised: 7th July 2025
Accepted: 24th September 2025
Available online: 26th January 2026

Keywords:

Air quality prediction;
Kernel Extreme Learning Machine;
Machine learning;
Regression technique.

ABSTRACT

Air pollution presents a substantial threat to human health, especially in urban areas like Jakarta, Indonesia, which ranked eleventh worldwide for poor air quality and urban pollution in mid-2025. This study is conducted with the objective of forecasting air quality over a designated future period by employing two advanced machine learning techniques: the Extreme Learning Machine (ELM) and its kernel-based variant, the Kernel Extreme Learning Machine (K-ELM). These methodologies are applied to predict the concentrations of five features of pollutants—PM10 (Particulate Matter), SO2 (Sulfur Dioxide), CO (Carbon Monoxide), O3 (Ozone), and NO2 (Nitrogen Dioxide)—which are critical indicators of environmental air quality and have significant implications for human health and environmental sustainability. Both methods are evaluated for their efficiency in time series regression, with a focus on training speed and generalization performance. The results demonstrate that the K-ELM model, especially when utilizing a Laplacian kernel, outperforms the standard ELM in predicting air quality based on the air quality index (AQI) dataset. Performance metrics indicate that K-ELM achieves superior accuracy, with an RMSE of 0.041, MSE of 0.002, MAE of 0.019, and an R-squared value of 0.898, confirming its effectiveness for air quality prediction in Jakarta. Furthermore, the Nemenyi post-hoc analysis across all metrics showed that K-ELM with the Laplacian kernel consistently achieved the highest rank and exhibited statistically significant improvements in multiple pairwise comparisons.

How to cite this article:

M. Kallista, I. P. D. Wibawa, and S. C. Obie, "EVALUATING THE EFFECTIVENESS OF KERNEL EXTREME LEARNING MACHINES OVER CONVENTIONAL ELM FOR AIR QUALITY INDEX PREDICTION", *BAREKENG: J. Math. & App.*, vol. 20, no. 2, pp. 1373-1388, Jun, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Air quality has emerged as a significant concern in modern society, driven by factors such as rapid urbanization, industrialization, and population growth [1], [2]. Poor air quality has significant implications for human health, particularly concerning respiratory [3], leading to increased morbidity and mortality. Studies also show that pollution is a major, yet often overlooked, risk factor for cardiovascular disease, accounting for over 60% of pollution-related deaths globally [4]. Additionally, it exerts adverse effects on environmental conditions and climate systems [5]. Accurate and timely air quality prediction is crucial for public health planning, environmental protection, and regulatory decision-making [6], [7]. Air quality prediction provides the estimation of concentration levels for multiple air pollutants, including particulate matter ($PM_{2.5}$, PM_{10}) [8], nitrogen dioxide (NO_2), sulfur dioxide (SO_2), carbon monoxide (CO), and ozone (O_3) [9], [10]. The concentration of these pollutants is affected by various dynamic factors, such as weather conditions, emissions from vehicles, industries, and changing seasons, making air quality forecasting a complex and challenging task [11]-[13].

The recent developments in data collection methods, including satellite observations, ground-based sensors, and IoT networks, alongside the development of machine learning techniques, have led to notable enhancements in the accuracy and scope of air quality prediction systems [14], [15]. According to IQAir, Jakarta was ranked eleventh among the world's most polluted cities in the world in May 2025. Its AQI (Air Quality Index) was 114, indicating unhealthy air quality, especially for vulnerable groups [16]. The policy note, developed by the Low Carbon Development Indonesia (LCDI) initiative under the Ministry of National Development Planning (Bappenas), highlights that air pollution from the transportation sector is a major contributor to premature deaths and respiratory illnesses in Indonesia. It calls for urgent action through cleaner transportation technologies and integrated urban planning to protect public health and support sustainable development [17].

Machine Learning (ML), a branch of Artificial Intelligence (AI) [18], has become an essential tool in the field of air quality prediction. By analyzing large datasets from environmental monitoring systems, meteorological records, and urban traffic patterns, ML algorithms can identify complex relationships and forecast pollutants such as particulate matter ($PM_{2.5}$, PM_{10}), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), carbon monoxide (CO), and ozone (O_3). These predictive capabilities are critical for issuing timely public health advisories, supporting evidence-based urban planning, and informing environmental policy decisions. Recent advancements in air quality prediction have significantly enhanced the ability to issue timely public health advisories, support urban planning, and inform environmental policy decisions. One notable development is AirNet, a machine learning-based model that forecasts air quality across more than 23,000 cities using algorithms like Random Forest and SVM. It provides real-time alerts through a web interface, helping mitigate health risks and guide environmental strategies [19]. Another cutting-edge approach is a 2025 study that combines Long Short-Term Memory (LSTM) networks with Transformer architectures to achieve high-precision urban air quality predictions. This hybrid model is particularly valuable for proactive health interventions and city planning [20].

This paper presents a comparison of Machine Learning methods for prediction, specifically focusing on the Extreme Learning Machine (ELM) and the Kernel Extreme Learning Machine (K-ELM). The Extreme Learning Machine is structured as a feed-forward artificial neural network featuring a single hidden layer, and it is also known as a single hidden layer feed-forward neural network (SLFN) [21]. ELM has demonstrated effective outcomes as a learning technique for addressing regression and classification problems [22]. Kernel Extreme Learning Machine (K-ELM) represents an advancement of the Extreme Learning Machine (ELM) methodology. Despite its high performance, ELM exhibits limitations in producing unstable prediction results [23]. The quantity of hidden neurons has a direct impact on the accuracy of outcomes. Research findings indicate that K-ELM demonstrates superior prediction accuracy compared to ELM [24]. K-ELM demonstrates the capability to address the variation issue arising from the random weights and biases established in ELM, while concurrently preserving the efficiency of the learning process [25].

The implementation of the ELM and KELM methods in the regression process aims to facilitate air quality predictions for a specified future timeframe. Recent investigations into air quality prediction demonstrate that Extreme Learning Machine (ELM) and its kernel variant (KELM) achieve notable generalization performance, offering advantages such as robust modeling [26], reliable forecasting [27], high computational efficiency [28], effective feature-based analysis [29], adaptive intelligent control strategies [30], and enhanced capabilities for environmental monitoring [31]. In this paper, we present a study aimed at

determining the accuracy of predicting air quality data in DKI Jakarta through the application of Extreme Learning Machine (ELM) and Kernel Extreme Learning Machine (K-ELM) methodologies. The evaluation results of Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-Squared are compared between the Extreme Learning Machine (ELM) and Kernel Extreme Learning Machine (K-ELM) algorithms. This research focuses on predicting air quality through the analysis of five particles in the air concentration, such as PM_{10} (Particulate Matter), SO_2 (Sulfur Dioxide), CO (Carbon Monoxide), O_3 (Ozone), and NO_2 (Nitrogen Dioxide). The dataset utilized in this research is the AQI dataset sourced from Jakarta Open Data, available as open source from the years 2017 to 2022. The dataset comprises 2.191 rows of data pertaining to five regions in DKI Jakarta: DKI1 (Bundaran HI), DKI2 (Kelapa Gading), DKI3 (Jagakarsa), DKI4 (Lubang Buaya), and DKI5 (Kebon Jeruk). The values of the five particles serve as features in the machine learning regression analysis conducted in this study. This study contributes by highlighting the comparative performance of different kernel functions in KELM against ELM for air quality prediction using Jakarta's AQI dataset, offering new insights into the most effective prediction modeling.

2. RESEARCH METHODS

The Extreme Learning Machine is structured as a feed-forward artificial neural network featuring a single hidden layer, commonly known as a single hidden layer feed-forward neural network (SLFN) [21]. In prior research, we developed a variant of ELM that emphasizes enhancing generalization performance through a control model approach [23], [32], [33]. The ELM algorithm demonstrates an outstanding learning speed during dataset training when compared to other algorithms, such as SVM and backpropagation neural networks [34].

2.1 Extreme Learning Machine

The Extreme Learning Machine (ELM) operates as a batch learning algorithm, meaning it processes the entire training dataset simultaneously in a single learning phase rather than incrementally. This approach offers several advantages that make ELM particularly suitable for large-scale regression and classification tasks. Notably, ELM is characterized by its exceptionally fast training speed, as it eliminates the need for iterative weight updates by randomly assigning input weights and biases and analytically determining output weights [35]. This simplicity in implementation reduces computational complexity and avoids issues such as local minima. Despite its straightforward architecture, ELM demonstrates strong generalization performance, especially when the number of hidden nodes is appropriately configured. Furthermore, its scalability and efficiency in handling large datasets make it a practical choice for real-time applications.

The ELM process is a batch learning process that learns all datasets in one go. Fig. 1 is the architecture of the Extreme Learning Machine (ELM). The output of a Single Layer Feedforward Neural Network (SLFN) with L hidden nodes, given an input vector $\mathbf{x_j}$ and N training samples, is computed as the weighted sum of the outputs of the hidden layer nodes. Each hidden node applies an activation function to a linear combination of the input vector and a bias term. The formulation can be expressed as:

$$f_L(\mathbf{x}_j) = \sum_{i}^{L} g(\mathbf{w}_i \cdot \mathbf{x}_j + b_i)\beta_i, \quad j = 1, \dots, N, \tag{1}$$

where $f_L(\mathbf{x}_j)$ is the predicted output for the input vector $\mathbf{x}_j$ from j-th training sample, L is the number of hidden nodes in the network, β_i is the output weight connecting the i-th hidden node to the output node, $g(\cdot)$ is the activation function (such as sigmoid, ReLU, or radial basis function), $\mathbf{w}_i$ is the weight vector connecting the input layer to the i-th hidden node, b_i is the bias term for the i-th hidden node, and $\mathbf{x}_j$ is the input vector. Eq. (1) can be written as $\mathbf{f}(\mathbf{x}) = \mathbf{h}(\mathbf{x})\boldsymbol{\beta}$, where $\mathbf{h}(\mathbf{x})$ denotes the hidden layer feature mapping, and $\boldsymbol{\beta} = [\beta_1 \quad \cdots \quad \beta_N]^T$ [36]. Given the corresponding target from training samples, $\mathbf{T} = [t_1 \quad \cdots \quad t_N]^T$, we have $\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$, where $\mathbf{H} = \begin{bmatrix} g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(\mathbf{w}_L \cdot \mathbf{x}_N + b_L) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_N + b_L) \end{bmatrix}$.

2.2 Kernel-Extreme Learning Machine

The flexibility of ELM also allows it to be adapted to various domains, including time-series forecasting, image recognition, and environmental modeling, particularly when extended through kernel-based methods such as Kernel ELM. In the newly developed ELM kernel, a positive regularization coefficient is introduced into the learning system to make it more stable [37]. Assuming $\mathbf{H^T H}$ is non-singular, the coefficient $\frac{\mathbf{I}}{C}$ is added to the diagonal of $\mathbf{H}^T\mathbf{H}$ in the calculation of the output weight $\boldsymbol{\beta}$. The resulting solution is more stable, and with better generalization performance, we can have $\boldsymbol{\beta} = \mathbf{H}^T\left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^T\right)^{-1}$.

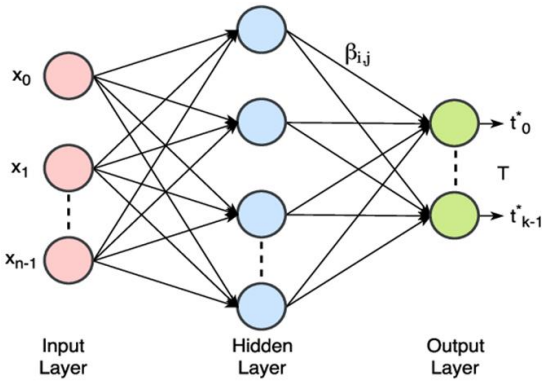

Figure 1. Architecture of Extreme Learning Machine (ELM)

The function corresponding to the regularized ELM is:

$$\mathbf{f(x)} = \mathbf{h(x)}\boldsymbol{\beta} = \mathbf{h(x)}\mathbf{H}^T\left(\frac{\boldsymbol{I}}{C} + \mathbf{H}\mathbf{H}^T\right)^{-1}\mathbf{T}, \tag{2}$$

where C is a regularization coefficient. It can be seen in ELM that the kernel matrix for ELM can be defined as follows. Suppose $\boldsymbol{\Omega}_{\mathbf{ELM}} = \mathbf{H}\mathbf{H}^{\mathbf{T}}, \boldsymbol{\Omega}_{\mathbf{ELM}_{\mathbf{i,j}}} = \mathbf{h}(\mathbf{x_i})\mathbf{h}(\mathbf{x_j}) = \mathbf{K}(\mathbf{x_i}, \mathbf{x_j})$. The ELM output function can be written concisely as:

$$\mathbf{f(x)} = \mathbf{h(x)}\mathbf{H}^T\left(\frac{\mathbf{I}}{C} + \mathbf{H}\mathbf{H}^{\mathbf{T}}\right)^{-1}\mathbf{T} = \begin{bmatrix}\mathbf{K(x, x_1)} \\ \vdots \\ \mathbf{K(x, x_N)}\end{bmatrix}^T\left(\frac{\mathbf{I}}{C} + \boldsymbol{\Omega}_{\mathbf{ELM}}\right)^{-1}\mathbf{T}. \tag{3}$$

In this implementation, the hidden layer feature mapping $\mathbf{h(x)}$ does not need to be known by the user; instead, the corresponding kernel $\mathbf{K(x, x_1)}$ can be calculated. The following is the architecture of the Kernel Extreme Learning Machine (K-ELM), as illustrated in Fig. 2.

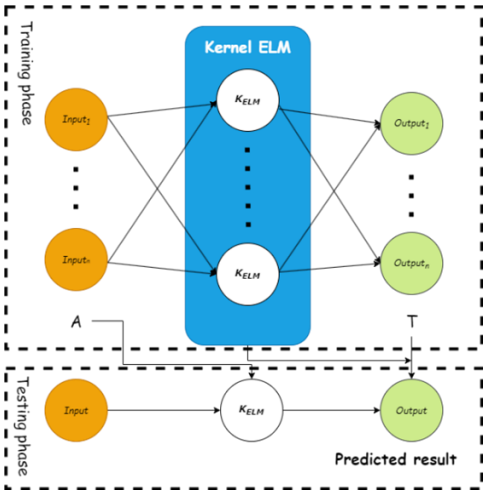

Figure 2. Architecture of Extreme Learning Machine (ELM)

The Kernel Extreme Learning Machine (K-ELM) is utilized to predict air quality indicators based on environmental data, leveraging its strong generalization capabilities on datasets. The methodology involves constructing a kernel matrix from the input data, selecting an appropriate kernel function along with its parameters, and solving a regularized least squares problem to determine the output weights. This method improves predictive accuracy while retaining the computational efficiency characteristic of the original ELM model.

2.3 Data Pre-processing

To perform regression using machine learning methods, it is essential to first examine the dataset. The AQI dataset used in this study contains missing values, which are addressed using the KNN Imputer from the scikit-learn library. This method is chosen for its effectiveness and accuracy compared to other imputation techniques. After the imputation process, the dataset is normalized using Min-Max Scaling to ensure compatibility with machine learning algorithms. This scaling enables effective prediction using the Extreme Learning Machine (ELM) and Kernel Extreme Learning Machine (K-ELM) method.

2.4 Utilizing Kernel

The kernel function can calculate the inner product directly on the higher-dimensional space (feature space) without the need to understand the mapping function from the input space to the feature space. In other words, the kernel function is the inner product function in the feature space. Choosing the right kernel will improve the performance of machine learning. RBF, Laplacian, Linear, Sigmoid, and Cosine kernels are some popular kernel functions. In this research, we further evaluate different types of kernel functions to examine their impact on the learning performance of the model. The selected kernel functions are as follows.

1. Radial Basic Function (RBF) Kernel

$$\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)_{\mathbf{rbf}} = \mathbf{exp}\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\mathbf{2\sigma^2}}\right), \tag{1}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are input vectors (data points), $\|\mathbf{x}_i - \mathbf{x}_j\|^2$ is the squared Euclidean distance between the two vectors, and sigma is a parameter that controls the width of the Gaussian function. The purpose of the RBF kernel is to measure similarity between points; closer points have higher values.

2. Laplacian Kernel

$$\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)_{\mathbf{lap}} = \mathbf{exp}(-\gamma|\mathbf{x}_i - \mathbf{x}_j|), \tag{2}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are input vectors (data points), $\|\mathbf{x}_i - \mathbf{x}_j\|$ is the Euclidean distance between the two vectors, γ is a kernel parameter that affects the decay rate of the exponential function. The purpose of the Laplacian kernel is similar to the RBF kernel but with a sharper drop-off in similarity.

3. Linear Kernel

$$\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)_{\mathbf{lin}} = (\mathbf{x}_i \cdot \mathbf{x}_j), \tag{3}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are input vectors (data points), $\mathbf{x}_i \cdot \mathbf{x}_j$ is the dot product of the two vectors. The purpose of the Linear kernel is to measure linear similarity between points.

4. Sigmoid Kernel

$$\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)_{\mathbf{sig}} = \mathbf{tanh}(\alpha(\mathbf{x}_i \cdot \mathbf{x}_j) + c), \tag{4}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are input vectors (data points), α is the slope parameter that scales the dot product, c is the offset or bias term, tanh is the hyperbolic tangent function. The purpose of the Sigmoid kernel is to mimic the behavior of neural networks and can be used for nonlinear classification.

5. Cosine Kernel

$$\mathbf{K}(\mathbf{x}_i, \mathbf{x}_j)_{\mathbf{cos}} = \frac{\mathbf{x}_i \cdot \mathbf{x}_j}{|\mathbf{x}_i||\mathbf{x}_j|}, \tag{5}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are input vectors (data points), $\mathbf{x}_i \cdot \mathbf{x}_j$ is the dot product of the two vectors, $\|\mathbf{x}_i\|$ and $\|\mathbf{x}_j\|$ are the norms (magnitudes) of the vectors. The purpose of the Cosine kernel is to measure the cosine of the angle between two vectors, which is useful when the magnitude is less important than the direction.

2.5 Training Process

The dataset in this study is partitioned using the k-Fold Cross-Validation method [38], as illustrated in Fig. 3. This technique involves randomly dividing the original dataset into k equally sized subsets or "folds." In this paper, the number of folds is set to five ($k = 5$), meaning the model is trained and evaluated five times, each time using a different fold as the test set and the remaining folds as the training set. This results in an 80:20 ratio between training and testing data in each iteration. The training process of the K-ELM method begins with kernel initialization, followed by the assignment of the regularization coefficient (C) and kernel parameter, which influence the shape and behavior of the kernel function. The kernel function is then computed to generate the training data's kernel matrix, denoted as $\mathbf{\Omega_{KELM}}$, as defined in Eq. (9).

1. Kernel initialization, coefficient values (C), and a kernel parameter that controls the shape of the kernel function.
2. Calculation of the kernel function to form the training data omega matrix (Ω_{KELM}). The matrix form can be seen in Eq. (9).

$$\mathbf{\Omega_{KELM}} = \begin{bmatrix} \mathbf{K}(\mathbf{x_1}, \mathbf{x_1}) & \cdots & \mathbf{K}(\mathbf{x_1}, \mathbf{x_N}) \\ \vdots & \ddots & \vdots \\ \mathbf{K}(\mathbf{x_N}, \mathbf{x_1}) & \cdots & \mathbf{K}(\mathbf{x_N}, \mathbf{x_N}) \end{bmatrix}_{N \times N}, \tag{6}$$

where N is the number of training data, $\mathbf{K}(\mathbf{x_i}, \mathbf{x_j})$ is the kernel function of i-th data and j-th data. This paper applies the RBF kernel function [39], the Laplacian kernel function [35], the Linear kernel function [40], Sigmoid, and Cosine kernel functions [41] formulated in Eqs. (4)-(8).

3. Calculate the output weight ($\boldsymbol{\beta}$) value with Eq. (10) as follows

$$\boldsymbol{\beta} = \left(\frac{\mathbf{I}}{C} + \mathbf{\Omega_{KELM}}\right)^{-1} \mathbf{T}, \tag{7}$$

where, $\mathbf{\Omega_{KELM}}$ is the omega matrix of the training data, $\left(\frac{\mathbf{I}}{C} + \mathbf{\Omega_{KELM}}\right)^{-1}$ is the inverse operation of the matrix, $\mathbf{I}$ is the identity matrix, C is the regularization of parameters or coefficients, and $\mathbf{T}$ is the target training data.

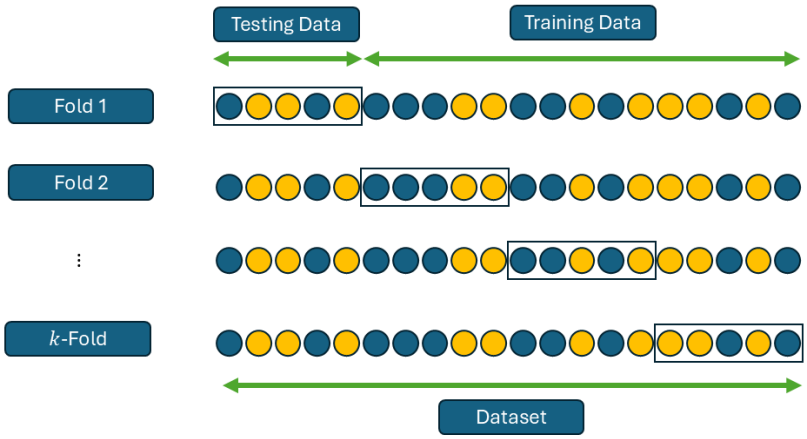

Figure 3. *k*-Fold Cross-Validation

2.6 Testing

The objective of this process is to evaluate the implementation of the K-ELM technique based on the training conducted in the previous stage. The following outlines the steps involved in the K-ELM testing procedure.

1. Retrieve the output weight value from the training process.
2. The calculation of the kernel function is to form the testing data omega matrix (Ω_{KELM}). The matrix form can be seen in Eq. (11).

$$\boldsymbol{\Omega}_{\mathbf{KELM}} = \begin{bmatrix} \mathbf{K}(\mathbf{x}, \mathbf{x_1}) \\ \vdots \\ \mathbf{K}(\mathbf{x}, \mathbf{x_N}) \end{bmatrix}^{T}_{N \times N}, \tag{8}$$

where, $K(x, x_1)$ is the kernel function of the j-th pair of testing data and training data $j = 1,2,..,N$.

3. Calculate the output target value using Eq. (12).

$$\hat{\boldsymbol{y}} = \boldsymbol{\Omega}_{\mathbf{KELM}}\, \boldsymbol{\beta}, \tag{9}$$

where, $\hat{\boldsymbol{y}}$ is the target output, $\boldsymbol{\Omega}_{\mathbf{KELM}}$ is the testing data omega matrix, $\boldsymbol{\beta}$ is the output weight matrix.

2.7 Denormalization

The denormalization process is intended to convert normalized values back to their original scale in the AQI dataset. This involves restoring the data using the original minimum and maximum values. The formula used for denormalization is presented in Eq. (13).

$$d = d'^{(max-min)} + min, \tag{10}$$

where, d is the normalized value of the prediction, d' is the denormalized value, min is the smallest value in each input feature, max is the largest value in each input feature.

2.8 Evaluation

Prediction performance can be evaluated using standard model assessment metrics. In this study, the evaluation is conducted using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the Coefficient of Determination (R^2). For MSE, RMSE, and MAE, lower values—ideally approaching zero—indicate better model performance. In contrast, for R^2, values closer to 1 signify a stronger correlation between predicted and actual values, reflecting higher predictive accuracy.

The pipeline process for the comparative study of KELM-based air quality prediction is shown in Fig. 4. The pipeline process begins with preprocessing, where the input data are normalized, and missing values are handled to ensure data quality. The dataset is then subjected to K-fold cross-validation during the splitting phase to improve model robustness and reduce overfitting. In the training stage, the Kernel Extreme Learning Machine (KELM) is applied with appropriate kernel selection and hyperparameter tuning. Finally, the model's performance is assessed in the evaluation stage using common metrics such as MAE, RMSE, and R^2.

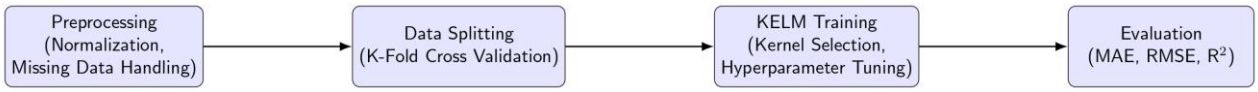

Figure 4. The Pipeline on Comparative Study of KELM-Based Air Quality Prediction

3. RESULTS AND DISCUSSION

This study presents the evaluation results of the prediction performance matrices for the Extreme Learning Machine (ELM) and Kernel Extreme Learning Machine (K-ELM) methods. In the ELM-based prediction algorithm, key parameters include the number of hidden neurons and the choice of activation functions. Meanwhile, the parameters used in the prediction algorithm using the K-ELM method are kernel,

regularization coefficient (C), and kernel parameter. These parameters affect the results of the machine learning model. This can be seen from the evaluation value of the matrix on the prediction results of Extreme Learning Machine (ELM) and Kernel Extreme Learning Machine (K-ELM). In addition to the kernel function, the Kernel Extreme Learning Machine (K-ELM) method also uses the kernel parameter and regularization coefficient (C). A kernel parameter is a parameter that controls the shape of the kernel function used in mapping data to a higher feature space. The kernel parameter values for the Laplacian kernel used in the prediction process using the Kernel Extreme Learning Machine are $\gamma = 0.0001$, $\gamma = 0.001$, and $\gamma = 0.01$. Regularization coefficient (C) is a parameter that can control the complexity of the model and overcome over-fitting. The regularization coefficient values (C) used in the prediction process using Kernel Extreme Learning Machine are 10, 20, and 30. The hyperparameters of the KELM model, including γ and the regularization coefficient (C), were tuned using a systematic grid search combined with cross-validation on the training dataset. All computations in this project were performed using Google Colaboratory with the Python programming language [42].

Fig. 5 illustrates PM_{10} air quality predictions from 2017 to 2023 generated using the Laplacian kernel, comparing actual and predicted values for both training and testing datasets. The black and magenta dots represent actual and predicted PM_{10} levels during training, respectively, while the blue and orange lines show actual and predicted values during testing. The close alignment between the predicted and actual values in both sets indicates that the model effectively learned the data. Furthermore, from Table 1, the learning process applied to this dataset is not well-suited for feature mapping using linear and cosine kernels, as these kernels are ineffective in separating the data. This inadequacy is evidenced by a negative R^2 value, which indicates poor model performance. The performance metrics for both the training and testing phases are presented in Table 1. Among the evaluated kernel functions, the Laplacian kernel demonstrates the best predictive performance, as reflected by the lowest error metrics. Fig. 5 presents the PM_{10} data plot for the DKI4 region, generated using the Laplacian kernel.

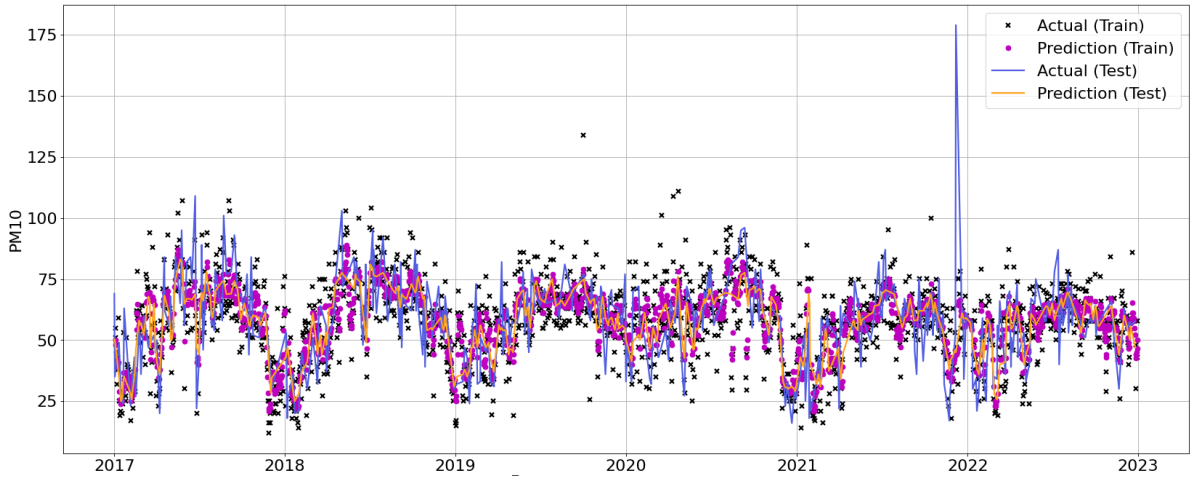

Figure 5. Plot of PM_{10} DKI4 K-ELM

Fig. 6 is the plot of SO_2 DKI3 with the Laplacian kernel. The time series plot presents the predicted and actual SO_2 concentrations from 2017 to 2023. Actual and predicted values for both training and testing phases are shown using distinct color-coded dots. The model effectively captures the overall trend and key fluctuations, including peaks in mid-2020 and early 2021. The close alignment between predicted and actual values indicates satisfactory model performance and generalization capability for SO_2 air quality forecasting. Fig. 7 illustrates the time series plot of carbon monoxide (CO) concentrations in the DKI3 region from 2017 to 2023, using the Laplacian kernel for prediction. The plot includes four data series: actual training data (black dots), predicted training data (magenta dots), actual test data (blue line), and predicted test data (orange line). The model demonstrates a strong ability to capture the temporal patterns of CO levels, with the predicted values closely aligning with the actual observations in both training and testing phases. This alignment indicates that the Laplacian kernel effectively models the underlying structure of the CO data, supporting its suitability for air quality forecasting in the DKI3 region.

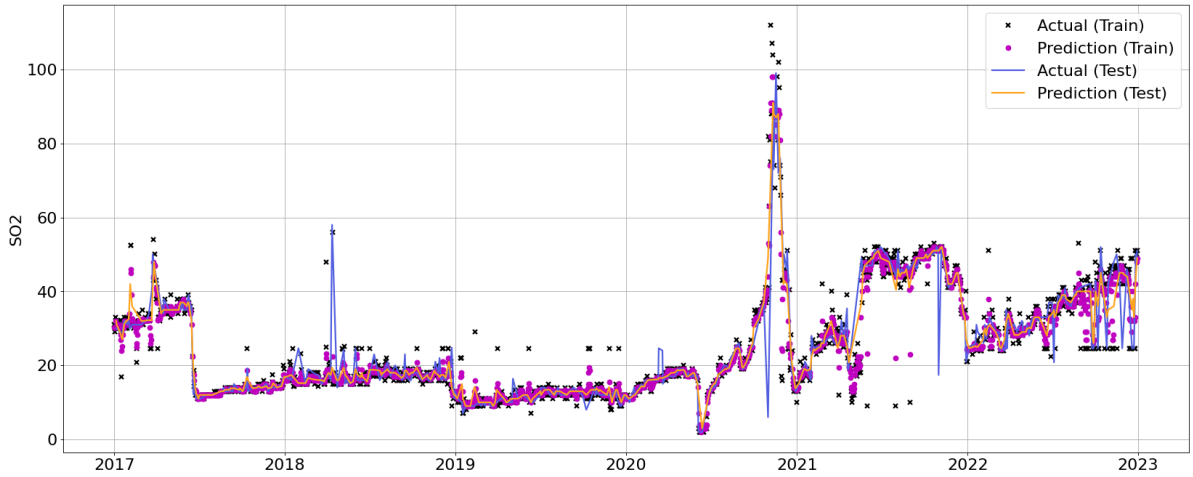

Figure 6. Plot of SO_2 DKI3 K-ELM

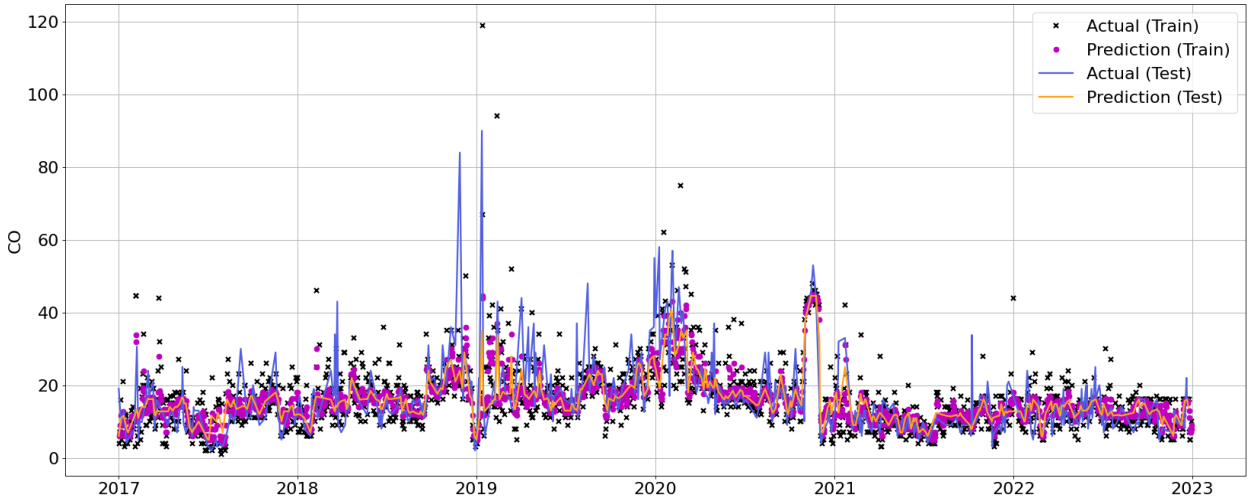

Figure 7. Plot of CO DKI3 K-ELM

Fig. 8 presents the time series plot of ozone (O_3) concentrations in the DKI1 region from 2017 to 2023, modeled using the Laplacian kernel. The plot includes actual and predicted values for both training and testing phases, with black and magenta dots representing actual and predicted training data, respectively, and blue and orange lines representing actual and predicted test data. The model effectively captures the temporal dynamics of O_3 levels, particularly the elevated concentrations observed between 2017 and early 2020. The close alignment between predicted and actual values across both phases indicates that the Laplacian kernel provides a reliable fit for forecasting O_3 concentrations in this region. Fig. 9 displays the time series plot of nitrogen dioxide (NO_2) concentrations in the DKI2 region from 2017 to 2023, modeled using the Laplacian kernel. The plot includes actual and predicted values for both training and testing phases: black and magenta dots represent actual and predicted training data, respectively, while blue and orange lines represent actual and predicted test data. The model demonstrates a strong ability to replicate the temporal patterns of NO_2 concentrations, with predicted values closely following the actual observations across both phases. This alignment indicates that the Laplacian kernel effectively captures the underlying structure of the NO_2 data, supporting its suitability for air quality prediction in the DKI2 region.

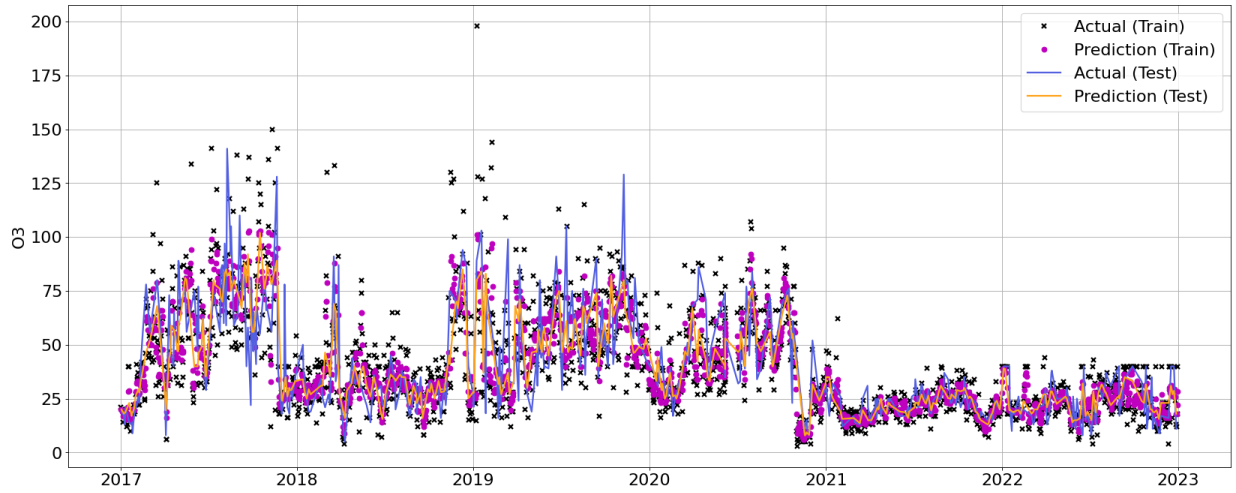

Figure 8. Plot of O_3 DKI1 K-ELM

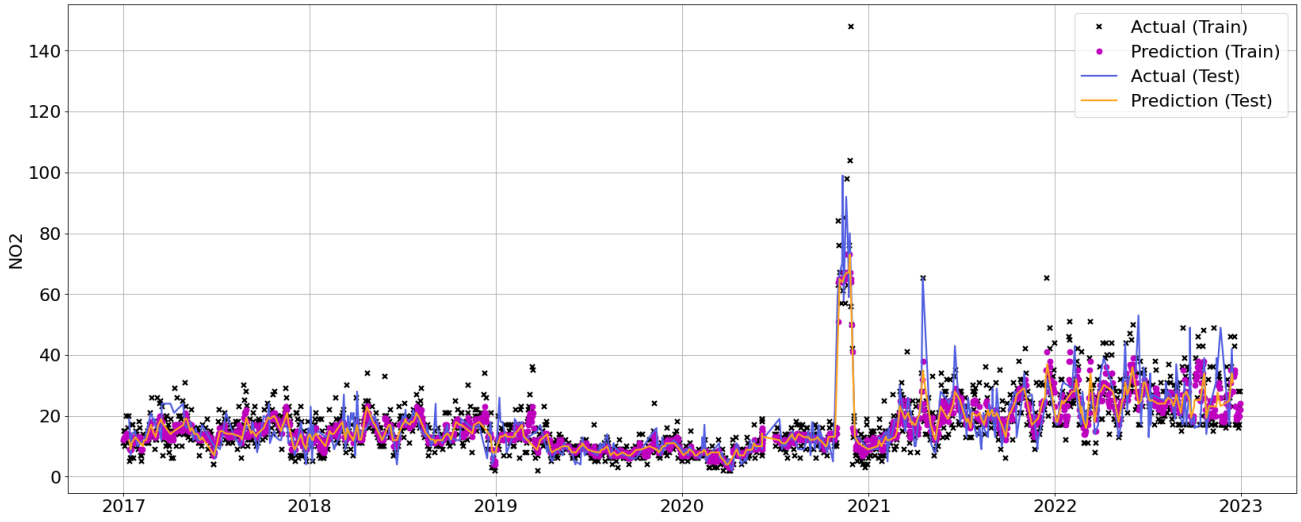

Figure 9. Plot of NO_2 DKI2 K-ELM

From Table 1, the performance evaluation of the Extreme Learning Machine (ELM) and Kernel Extreme Learning Machine (K-ELM) methods for air quality prediction in Jakarta reveals that K-ELM consistently outperforms ELM across all gas types—PM_{10}, O_3, SO_2, NO_2, and CO. Among the various kernel functions tested, the Laplacian kernel yields the most accurate results, achieving the lowest RMSE, MSE, and MAE values, along with the highest R^2 scores in both training and testing phases. This indicates its superior ability to model nonlinear relationships inherent in air quality data. In contrast, linear and cosine kernels generally perform poorly, often resulting in negative or undefined R^2 values, which reflect inadequate model fit. For each pollutant, the Laplacian kernel within the K-ELM framework consistently provides the best predictive performance, with R^2 values reaching as high as 0.898 for SO_2 and 0.700 for NO_2 in the testing phase. These findings underscore the effectiveness of kernel-based learning, particularly with the Laplacian kernel, in enhancing the accuracy and reliability of air quality forecasting models.

Table 1. Performance Evaluation on the Regression Task of Air Quality Prediction in Jakarta

Method	Evaluation of Training Data				Evaluation of Testing Data				Kernel type	Gas type
	RMSE	MSE	MAE	R^2	RMSE	MSE	MAE	R^2		
ELM	0.07	0.005	0.054	0.565	0.076	0.006	0.055	0.142	-	PM_{10}
K-ELM	**0.045**	**0.002**	**0.034**	**0.788**	**0.071**	**0.005**	**0.053**	**0.488**	Laplacian	
	0.062	0.004	0.045	0.621	0.075	0.006	0.056	0.43	RBF	
	0.101	0.01	0.079	-	0.102	0.01	0.079	-	Linear	
	0.097	0.009	0.074	0.062	0.099	0.01	0.076	0.016	Sigmoid	
	0.099	0.01	0.074	-	0.102	0.01	0.076	-	Cosine	
ELM	0.084	0.007	0.064	0.678	0.077	0.006	0.057	0.161	-	O_3

Method	Evaluation of Training Data				Evaluation of Testing Data				Kernel type	Gas type
	RMSE	MSE	MAE	R^2	RMSE	MSE	MAE	R^2		
	0.047	**0.002**	**0.031**	**0.861**	**0.079**	**0.006**	**0.051**	**0.602**	Laplacian	
	0.066	0.004	0.044	0.713	0.083	0.007	0.056	0.556	RBF	
K-ELM	0.114	0.013	0.83	0.151	0.117	0.014	0.084	0.133	Linear	
	0.109	0.012	0.076	0.245	0.11	0.012	0.077	0.225	Sigmoid	
	0.13	0.017	0.093	-	0.131	0.017	0.094	-	Cosine	
ELM	0.049	0.002	0.033	0.874	0.083	0.007	0.057	0.506	-	
	0.022	**0.001**	**0.011**	**0.969**	**0.041**	**0.002**	**0.019**	**0.898**	Laplacian	
	0.036	0.001	0.017	0.923	0.043	0.002	0.021	0.889	RBF	SO_2
K-ELM	0.145	0.021	0.119	-	0.149	0.022	0.123	-	Linear	
	0.097	0.009	0.06	0.441	0.097	0.009	0.059	0.44	Sigmoid	
	0.137	0.019	0.098	-	0.138	0.019	0.098	-	Cosine	
ELM	0.034	0.001	0.024	0.898	0.093	0.009	0.068	0.521	-	
	0.027	**0.001**	**0.018**	**0.869**	**0.041**	**0.002**	**0.027**	**0.7**	Laplacian	
	0.034	0.001	0.023	0.776	0.043	0.002	0.028	0.676	RBF	NO_2
K-ELM	0.091	0.008	0.075	-	0.102	0.01	0.083	-	Linear	
	0.068	0.005	0.038	0.188	0.068	0.005	0.039	0.187	Sigmoid	
	0.072	0.005	0.047	-	0.078	0.006	0.049	-	Cosine	
ELM	0.058	0.002	0.042	**0.777**	**0.039**	**0.002**	**0.03**	0.19	-	
	0.04	**0.002**	**0.024**	0.75	0.058	0.003	0.037	**0.465**	Laplacian	
	0.049	0.002	0.032	0.592	0.06	0.004	0.038	0.422	RBF	CO
K-ELM	0.08	0.006	0.054	-	0.081	0.007	0.055	-	Linear	
	0.074	0.006	0.047	0.115	0.074	0.006	0.048	0.106	Sigmoid	
	0.077	0.006	0.052	-	0.08	0.006	0.053	-	Cosine	

For R^2, the χ^2 value is 14.04 with a p-value of 0.003, also suggesting a significant difference in model performance. These findings confirm that the choice of kernel function has a meaningful impact on the accuracy and reliability of air quality predictions, justifying the use of statistical testing to support model selection. We use the significance level threshold (denoted as $\alpha = 0.05$) to determine whether the observed differences in model performance across kernel methods are statistically significant. From the test results, all four evaluation metrics—RMSE ($p = 0.001$), MSE ($p = 0.002$), MAE ($p = 0.001$), and R^2 ($p = 0.003$)—have p-values less than 0.05. This means that the null hypothesis (which assumes no difference in performance among the methods) can be rejected for each metric.

Table 2. Friedman's ANOVA Test on Air Quality Prediction in Jakarta

Effects	SS	df	MS	χ^2	$p > \chi^2$	Evaluation
Variety	69.1	5	13.82	19.74	0.001	
Error	18.4	20	0.92	-	-	Fr-Test RMSE
Total	87.5	29	-	-	-	
Variety	65.9	5	13.18	18.83	0.002	
Error	21.6	20	1.08	-	-	Fr-Test MSE
Total	87.5	29	-	-	-	
Variety	73.1	5	14.62	20.89	0.001	
Error	14.4	20	0.72	-	-	Fr-Test MAE
Total	87.5	29	-	-	-	
Variety	23.4	3	7.8	14.04	0.003	Fr-Test R^2

Effects	SS	df	MS	χ^2	$p > \chi^2$	Evaluation
Error	1.6	12	0.13	-	-	
Total	25	19	-	-	-	

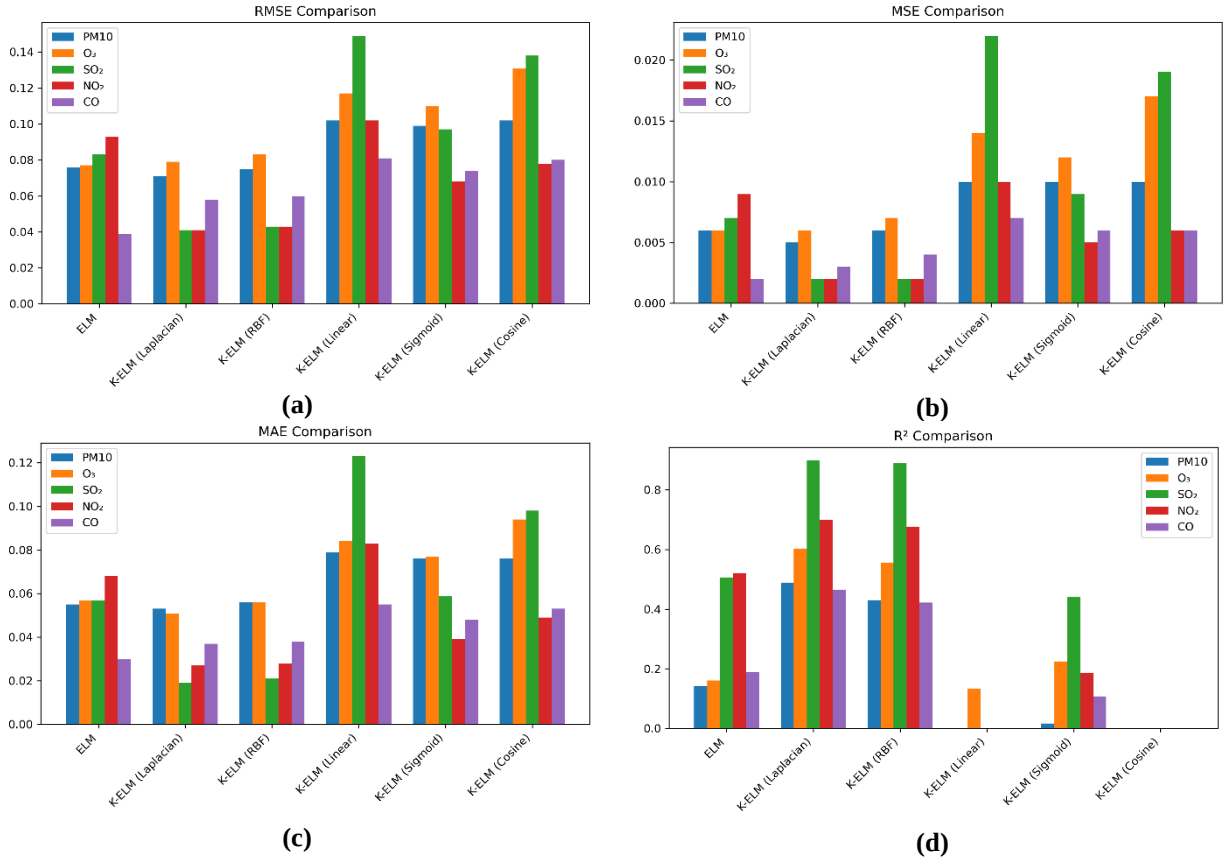

Figure 10. Performance Comparison of ELM vs. K-ELM Models for Air Quality Prediction in Jakarta (a) RMSE comparison, (b) MSE comparison, (c) MAE comparison, and (d) R^2 comparison

From Table 3, Friedman's mean rank analysis was utilized to compare the performance of different kernel methods across four evaluation metrics: RMSE, MSE, MAE, and R^2. Friedman's mean rank is a non-parametric statistical method used when the data does not meet the assumptions required for parametric tests, such as normal distribution or equal variances. From the results, the Laplacian kernel consistently achieved the lowest mean ranks across all metrics (RMSE: 1.4, MSE: 1.4, MAE: 1.2, R^2: 1.0), indicating superior predictive accuracy and model consistency. This suggests that the Laplacian kernel is the most effective among the tested methods for capturing the underlying patterns in air quality data. In contrast, the Linear and Cosine kernels exhibited the highest mean ranks (e.g., Linear RMSE: 5.6, Cosine RMSE: 5.2), reflecting comparatively poor performance. The ELM method, which does not utilize kernel mapping, showed moderate performance with mean ranks ranging from 2.4 to 3.2. These findings imply that kernel selection significantly influences model performance, with the Laplacian kernel offering the most robust results for regression tasks in air quality forecasting. The comparative performance analysis of ELM and Kernel-ELM models across multiple air pollutants using RMSE, MSE, MAE, and R^2 metrics is shown in Fig. 10.

Table 3. Friedman Mean Rank on Air Quality Prediction in Jakarta

Methods	Fr-Test RMSE	Fr-Test MSE	Fr-Test MAE	Fr-Test R^2	Kernel
ELM	2.6	2.4	2.8	3.2	-
K-ELM	**1.4**	**1.4**	**1.2**	**1**	**Laplacian**
	2.4	2.6	2.4	2	RBF
	5.6	5.4	5.8	-	Linear
	3.8	4	3.8	3.8	Sigmoid
	5.2	5.2	5	-	Cosine

Another tool for comparing multiple algorithms in non-parametric tests is the Nemenyi test. Nemenyi test is a post-hoc statistical test that indicates that there are differences among multiple groups, specifically to find which pairs of groups (or algorithms, models, treatments, etc.) are significantly different from each other. The critical difference (CD) in the Nemenyi test is a threshold for determining if the difference between two groups' average ranks is statistically significant, calculated from the number of groups, datasets, and significance level. The critical difference can be calculated using this formula:

$$CD = q_\alpha \sqrt{\frac{k(k+1)}{6N}}, \tag{11}$$

where k is number of algorithms, N is number of samples. For $\alpha = 0.05$ and $k = 6$, we have $q_\alpha \approx 2.850$, and for $k = 4$, we have $q_\alpha \approx 2.569$.

Table 4. Nemenyi Test on Air Quality Prediction in Jakarta

Algorithm	Nemenyi Test (Average Ranks)			
	RMSE	MSE	MAE	R^2
ELM	2.6	2.6	2.8	3.2
K-ELM (Laplacian)	1.4	1.5	1.2	1
K-ELM (RBF)	2.4	2.3	2.4	2
K-ELM (Linear)	5.7	5.6	5.8	-
K-ELM (Sigmoid)	3.8	4.1	3.9	3.8
K-ELM (Cosine)	5.1	4.9	4.9	-

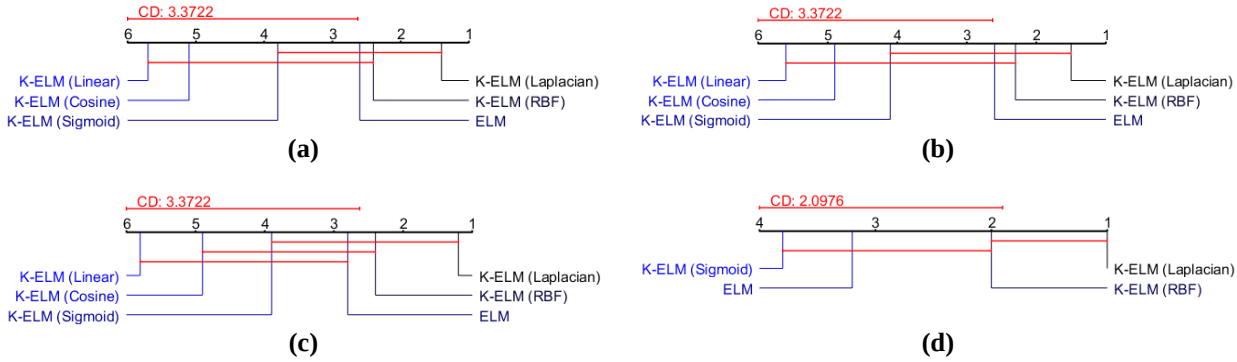

Figure 11. Nemenyi Test Diagram on Evaluation Performance (a). RMSE ($CD = 3.3722, \alpha = 0.05$), (b). MSE ($CD = 3.3722, \alpha = 0.05$), (c). MAE ($CD = 3.3722, \alpha = 0.05$) and (d). R^2 ($CD = 2.0976, \alpha = 0.05$)

For the RMSE evaluation, based on the numerical results in Table 4 and Fig.11 (a), the Nemenyi post-hoc test indicated that K-ELM with the Laplacian kernel achieved the best average rank, followed by K-ELM with the RBF kernel and the standard ELM method. With a Critical Difference (CD) of 3.3722, statistically significant differences were observed between K-ELM (Laplacian) and K-ELM (Cosine) (rank gap ≈ 3.5) and K-ELM (Linear) (rank gap ≈ 4.5), as well as between K-ELM (RBF) and K-ELM (Linear) (rank gap ≈ 3.8). For the MSE evaluation, based on the numerical results in Table 4 and Fig. 11 (b), the Nemenyi post-hoc test with a Critical Difference (CD) of 3.3722 revealed that only three pairwise differences were statistically significant: K-ELM (Laplacian) outperformed K-ELM (Cosine) (rank gap ≈ 3.5 > CD) and K-ELM (Linear) (rank gap ≈ 4.5 > CD), while K-ELM (RBF) outperformed K-ELM (Linear) (rank gap ≈ 3.8 > CD). All other rank gaps were below the CD threshold and therefore not statistically significant. For the MAE evaluation, based on the numerical results in Table 4 and Fig. 11 (c), the Nemenyi post-hoc test with a Critical Difference (CD) of 3.3722 revealed that only three pairwise differences were statistically significant: K-ELM (Laplacian) outperformed K-ELM (Cosine) (rank gap ≈ 3.5 > CD) and K-ELM (Linear) (rank gap ≈ 4.5 > CD), while K-ELM (RBF) outperformed K-ELM (Linear) (rank gap ≈ 3.8 > CD). For the R^2 evaluation, based on the numerical results in Table 4 and Fig. 11 (d), the Nemenyi post-hoc test with a Critical Difference (CD) of 2.0976 revealed that only the comparisons involving the top method were statistically significant: K-ELM (Laplacian) significantly outperformed K-ELM (Sigmoid) (rank gap ≈ 3.0 > CD) and ELM (rank gap

$\approx 2.1 > CD$). These rank gaps exceeded the CD threshold, confirming the statistical superiority of the Laplacian kernel in these pairwise comparisons.

4. CONCLUSION

The study evaluated Extreme Learning Machine (ELM) and Kernel-based ELM (K-ELM) models for predicting air quality across five key pollutants (PM10, O_3, SO_2, NO_2, and CO) using four regression metrics: RMSE, MSE, MAE, and R^2. The following summarizes the conclusions of the research conducted:

1. Among all configurations, K-ELM with the Laplacian kernel consistently outperformed other models, demonstrating superior accuracy and generalization across pollutants.
2. The standard ELM served as a reliable baseline but was generally surpassed by the kernelized variants, especially in capturing complex pollutant behaviors.
3. Specifically, the K-ELM model with a Laplacian kernel provided the most accurate predictions for Jakarta's air quality, outperforming both the standard ELM and other kernel configurations.
4. This study is limited to data from 2017–2022 with only five pollutant features, and the findings may not fully generalize to other regions or conditions, highlighting the need for broader datasets in future research. Numerical results on the DKI Jakarta AQI dataset (2017–2022) showed:
   a. ELM: RMSE of 0.083, MSE of 0.007, MAE of 0.057, R^2 of 0.506.
   b. K-ELM (Laplacian kernel): RMSE of 0.041, MSE of 0.002, MAE of 0.019, R^2 of 0.898.
5. The Laplacian kernel consistently ranked highest across all evaluation metrics in the Friedman test, with scores indicating its superior performance (1.4 for RMSE and MSE, 1.2 for MAE, and 1.0 for R^2).
6. Across all evaluation metrics (RMSE, MSE, MAE, and R^2), the Nemenyi post-hoc tests confirmed that K-ELM with the Laplacian kernel consistently achieved the best average rank and demonstrated statistically significant superiority in several pairwise comparisons.
7. By comparing ELM and KELM, we provide insights into the selection of efficient and reliable predictive models that have the potential to strengthen early warning systems and support targeted interventions aimed at reducing pollution levels. Air quality prediction performance and model accuracy can be improved through the integration of meteorological data, the use of hybrid models such as LSTM–KELM, and real-time deployment in future work.

Author Contributions

Meta Kallista: Conceptualization, Formal Analysis, Methodology, Supervision, Validation, Writing-Review and Editing. Ig. Prasetya Dwi Wibawa: Data Curation, Draft Preparation, Formal Analysis, Validation, Writing-Review and Editing. Sultan Chisson Obie: Data Curation, Visualization, Writing-Original Draft. All authors discussed the results and contributed to the final manuscript.

Funding Statement

This research received a grant from Telkom University for providing support through Skema Penelitian Mandiri, Universitas Telkom, number 778/LIT06/PPM-LIT/2025.

Acknowledgment

The authors would like to thank Telkom University for providing support through research funding. The authors also acknowledge the contributions of the Human Centric (HUMIC) and the Smart Transportation and Robotics (STAR), Center of Excellence, Telkom University, for valuable support and collaboration throughout this research.

Declarations

The authors declare no competing interests.

Declaration of Generative AI and AI-assisted Technologies

Generative AI tools (e.g., ChatGPT) were used solely for language refinement, including grammar, spelling, and clarity. The scientific content, analysis, interpretation, and conclusions were developed entirely by the authors. All final text was reviewed and approved by the authors.

REFERENCES

[1] V. Saxena, "WATER QUALITY, AIR POLLUTION, AND CLIMATE CHANGE: INVESTIGATING THE ENVIRONMENTAL IMPACTS OF INDUSTRIALIZATION AND URBANIZATION," *Water Air Soil Pollut*, vol. 236, no. 2, pp. 1–40, 2025. doi: https://doi.org/10.1007/s11270-024-07702-4

[2] N. Thanvisitthpon, K. Kallawicha, and H. J. Chao, "EFFECTS OF URBANIZATION AND INDUSTRIALIZATION ON AIR QUALITY," In *Health and Environmental Effects Of Ambient Air Pollution*, Elsevier, 2024, pp. 231–255. doi: https://doi.org/10.1016/B978-0-443-16088-2.00003-X

[3] R. A. Tavella *et al.*, "AN EXPLORATORY STUDY ON THE ASSOCIATION BETWEEN AIR POLLUTION AND HEALTH PROBLEMS (ICD-10) WITH AN EMPHASIS ON RESPIRATORY DISEASES," *Atmos Pollut Res*, vol. 16, no. 2, p. 102377, 2025. doi: https://doi.org/10.1016/j.apr.2024.102377

[4] R. A. Montone *et al.*, "IMPACT OF AIR POLLUTION ON ISCHEMIC HEART DISEASE: EVIDENCE, MECHANISMS, CLINICAL PERSPECTIVES," *Atherosclerosis*, vol. 366, pp. 22–31, 2023. doi: https://doi.org/10.1016/j.atherosclerosis.2023.01.013

[5] H. Cao *et al.*, "AIR POLLUTION, TEMPERATURE AND MUMPS: A TIME-SERIES STUDY OF INDEPENDENT AND INTERACTION EFFECTS," *Ecotoxicol Environ Saf*, vol. 291, p. 117826, 2025. doi: https://doi.org/10.1016/j.ecoenv.2025.117826

[6] S. Kant, "FROM DATA TO DECISION-MAKING: UTILIZING DECISION TREE FOR AIR QUALITY MONITORING IN SMART URBAN AREAS," *International Journal of Information Technology*, vol. 17, no. 1, pp. 665–672, 2025. doi: https://doi.org/10.1007/s41870-024-02208-y

[7] S. Jayaraman and A. S, "ENHANCING URBAN AIR QUALITY PREDICTION USING TIME-BASED-SPATIAL FORECASTING FRAMEWORK," *Sci Rep*, vol. 15, no. 1, p. 4139, 2025. doi: https://doi.org/10.1038/s41598-024-83248-z

[8] G. Ulpiani, E. Pisoni, J. Bastos, F. Monforti-Ferrario, and N. Vetters, "ARE CITIES READY TO SYNERGISE CLIMATE NEUTRALITY AND AIR QUALITY EFFORTS?," *Sustain Cities Soc*, vol. 118, p. 106059, 2025. doi: https://doi.org/10.1016/j.scs.2024.106059

[9] A. K. Rad, M. J. Nematollahi, A. Pak, and M. Mahmoudi, "PREDICTIVE MODELING OF AIR QUALITY IN THE TEHRAN MEGACITY VIA DEEP LEARNING TECHNIQUES," *Sci Rep*, vol. 15, no. 1, p. 1367, 2025. doi: https://doi.org/10.1038/s41598-024-84550-6

[10] S. Neelamegam, "ENHANCING ENVIRONMENTAL HEALTH THROUGH POLLUTANT DATA PREDICTION WITH GRADIENT BOOSTING," in *Advances in Electrical and Computer Technologies*, CRC Press, 2025, pp. 18–23. doi: https://doi.org/10.1201/9781003515470

[11] S. Mahato, S. Kundu, J. Cermak, and P. K. Joshi, "METEOROLOGICAL INFLUENCES ON AIR POLLUTION DYNAMICS IN POLLUTION EPICENTRE OF NATIONAL CAPITAL REGION, INDIA," *Chemosphere*, vol. 377, p. 144353, 2025. doi: https://doi.org/10.1016/j.chemosphere.2025.144353

[12] M. Nasar-u-Minallah, M. Jabeen, N. Parveen, M. Abdullah, and M. H. F. Nuskiya, "EXPLORING THE SEASONAL VARIABILITY AND NEXUS BETWEEN URBAN AIR POLLUTION AND URBAN HEAT ISLANDS IN LAHORE, PAKISTAN," *Acta Geophysica*, pp. 1–21, 2025. doi: https://doi.org/10.1007/s11600-025-01574-w

[13] M. A. Goheer, S. S. Hassan, A. S. Sheikh, Y. Malik, M. Uzair, and T. N. Satti, "ASSESSING SMOG TRENDS AND SOURCES OF AIR POLLUTANTS ACROSS NORTHEASTERN DISTRICTS OF PUNJAB, PAKISTAN USING GEOSPATIAL TECHNIQUES," *International Journal of Environmental Science and Technology*, vol. 22, no. 5, pp. 3657–3674, 2025. doi: https://doi.org/10.1007/s13762-024-05754-x

[14] S. Kumari *et al.*, "NEXT-GENERATION AIR QUALITY MANAGEMENT: UNVEILING ADVANCED TECHNIQUES FOR MONITORING AND CONTROLLING POLLUTION," *Aerosol Science and Engineering*, pp. 1–22, 2025. doi: https://doi.org/10.1007/s41810-024-00281-1

[15] A. M. Mutawa, A. AlShaibani, and L. A. Almatar, "A COMPREHENSIVE REVIEW OF DUST STORM DETECTION AND PREDICTION TECHNIQUES: LEVERAGING SATELLITE DATA, GROUND OBSERVATIONS, AND MACHINE LEARNING," *IEEE Access*, 2025. doi: 10.1109/ACCESS.2025.3541075

[16] Radio Republik Indonesia, "JAKARTA'S AIR QUALITY UNHEALTHY TODAY, AMONG WORLD'S WORST," 2025. Accessed: May 29, 2025. [Online]. Available: https://rri.co.id/en/archipelago/1522329/jakarta-s-air-quality-unhealthy-today-among-world-s-worst

[17] Bappenas, "THE IMPACT OF AIR POLLUTION FROM THE TRANSPORTATION SECTOR ON HEALTH IN INDONESIA," Oct. 2023. [Online]. Available: https://lcdi-indonesia.id/wp-content/uploads/2024/03/2-EN-Policy-Note-2023.10.16.pdf

[18] J. Jiang and S. Chen, "INFLUENCE OF ARTIFICIAL INTELLIGENT IN INDUSTRIAL ECONOMIC SUSTAINABILITY DEVELOPMENT PROBLEMS AND COUNTERMEASURES," *Heliyon*, vol. 10, no. 3, 2024. doi: https://doi.org/10.1016/j.heliyon.2024.e25079

[19] M. Rahman *et al.*, "AIRNET: PREDICTIVE MACHINE LEARNING MODEL FOR AIR QUALITY FORECASTING USING WEB INTERFACE," *Environmental Systems Research*, vol. 13, no. 44, 2024. doi: https://doi.org/10.1186/s40068-024-00378-z

[20] Y. Liu, M. Tee, L. Lu, F. Zhou, and B. Lu, "HIGH-PRECISION URBAN AIR QUALITY PREDICTION USING A LSTM-TRANSFORMER HYBRID ARCHITECTURE," *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 4, 2025. doi: https://doi.org/10.14569/IJACSA.2025.0160431

[21] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "EXTREME LEARNING MACHINE FOR REGRESSION AND MULTICLASS CLASSIFICATION," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 2, pp. 513–529, 2011. doi: https://doi.org/10.1109/TSMCB.2011.2168604

[22] G.-B. Huang, D. H. Wang, and Y. Lan, "EXTREME LEARNING MACHINES: A SURVEY," *International journal of machine learning and cybernetics*, vol. 2, pp. 107–122, 2011. doi: https://doi.org/10.1007/s13042-011-0019-y

[23] W. M. Sukiman, M. Kallista, Ig. P. Dwi Wibawa, and P. D. Kusuma, "CASE STUDY: JAKARTA AQI CLASSIFICATION USING EXTREME LEARNING MACHINE METHOD WITH OVERSAMPLING TECHNIQUE," in *Asia Pacific Conference on Wireless and Mobile*, Denpasar, Indonesia: IEEE, Oct. 2023, pp. 225–231. doi: https://doi.org/10.1109/APWiMob59963.2023.10365645

[24] X. Li *et al.*, "EMPIRICAL ANALYSIS: STOCK MARKET PREDICTION VIA EXTREME LEARNING MACHINE," *Neural Comput Appl*, vol. 27, pp. 67–78, 2016. doi: https://doi.org/10.1007/s00521-014-1550-z

[25] M. A. P. Nasution, I. Cholissodin, and Indriati, "PREDIKSI PRICE EARNING RATIO SAHAM MENGGUNAKAN ALGORITME KERNEL EXTREME LEARNING MACHINE (STUDI KASUS: PT TELKOM)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, vol. 4, pp. 3455–3462, 2020.

[26] M. V. Saikumar, F. Noorbasha, K. S. Rao, and K. G. Sravani, "AIR QUALITY INDEX PREDICTION USING AN ENHANCED EXTREME LEARNING MACHINE BASED ON GENETIC ALGORITHMS," *Int J Environ Sci*, vol. 11, no. 10s, pp. 21–30, 2025. doi: https://doi.org/10.64252/wcvfx368

[27] A. Binbusayyis, M. A. Khan, M. M. Ahmed A, and W. R. S. Emmanuel, "A DEEP LEARNING APPROACH FOR PREDICTION OF AIR QUALITY INDEX IN SMART CITY," *Discover Sustainability*, vol. 5, no. 1, p. 89, 2024. doi: https://doi.org/10.1007/s43621-024-00272-9

[28] B. Adepu and T. Archana, "FORECASTING AIR QUALITY INDEX USING IMPROVED EXTREME LEARNING MACHINE BASED ON GENETIC ALGORITHM," *Grenze International Journal of Engineering & Technology (GIJET)*, vol. 11, 2025.

[29] M. Lakshmipathy, S. B. Panjagal, M. J. Shanthi Prasad, and G. N. Kodandaramaiah, "HEALTH AND ECOLOGICAL RISK ASSESSMENT-BASED AIR QUALITY PREDICTION FRAMEWORK USING ENSEMBLE LEARNING NETWORK WITH OPTIMAL WEIGHTED PREDICTION SCORE," *Int J Image Graph*, p. 2750060, 2025. doi: https://doi.org/10.1142/S0219467827500604

[30] K. Agarwal, S. Chennupalli, R. Kolluru, and H. Vaibhav, "PREDICTIVE MODELING OF AIR QUALITY: PERFORMANCE COMPARISON OF SUPERVISED MACHINE LEARNING, DEEP LEARNING, AND EXTREME LEARNING MACHINES," in *2025 International Conference on Intelligent Computing and Control Systems (ICICCS)*, IEEE, 2025, pp. 1262–1269. doi: https://doi.org/10.1109/ICICCS65191.2025.10984959

[31] M. Wang, X. Lei, and X. Zhang, "RESEARCH ON CARBON EMISSION PREDICTION OF THE TRANSPORTATION INDUSTRY IN SHAANXI PROVINCE BASED ON DFE-IPOA-KELM MODEL," *Pol J Environ Stud*, vol. 34, no. 4, 2025. doi: https://doi.org/10.15244/pjoes/189718

[32] I. P. D. Wibawa, C. Machbub, A. S. Rohman, and E. Hidayat, "REDUCED COMPUTATION FOR EXTREME LEARNING MACHINE BASED ON CHOLESKY FACTORIZATION," in *2022 13th Asian Control Conference (ASCC)*, IEEE, 2022, pp. 488–493. doi: https://doi.org/10.23919/ASCC56756.2022.9828328

[33] I. P. D. Wibawa, C. Machbub, A. S. Rohman, and E. Hidayat, "MODIFIED ONLINE SEQUENTIAL EXTREME LEARNING MACHINE ALGORITHM USING MODEL PREDICTIVE CONTROL APPROACH," *Intelligent Systems with Applications*, vol. 18, p. 200191, 2023. doi: https://doi.org/10.1016/j.iswa.2023.200191

[34] S. Ding, X. Xu, and R. Nie, "EXTREME LEARNING MACHINE AND ITS APPLICATIONS," *Neural Comput Appl*, vol. 25, pp. 549–556, 2014. doi: https://doi.org/10.1007/s00521-013-1522-8

[35] J. Wang, S. Lu, S. H. Wang, and Y. D. Zhang, "A REVIEW ON EXTREME LEARNING MACHINE," *Multimed Tools Appl*, vol. 81, no. 29, pp. 41611–41660, 2022. doi: https://doi.org/10.1007/s11042-021-11007-7

[36] W. Huang, N. Li, Z. Lin, G.B. Huang, W. Zong, and J. Zhou, "LIVER TUMOR DETECTION AND SEGMENTATION USING KERNEL-BASED EXTREME LEARNING MACHINE," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan: IEEE, Jul. 2013, pp. 3662–3665. doi: https://doi.org/10.1109/embc.2013.6610337

[37] S. Li, G. Liu, and S. Xiao, "EXTREME LEARNING MACHINE WITH KERNELS FOR SOLVING ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS," *Cognit Comput*, vol. 15, p. 200191, 2023. doi: https://doi.org/10.1007/s12559-022-10026-2

[38] I. Gazalba *et al.*, "COMPARATIVE ANALYSIS OF K-NEAREST NEIGHBOR AND MODIFIED K-NEAREST NEIGHBOR ALGORITHM FOR DATA CLASSIFICATION," in *2017 2nd international conferences on information technology, information systems and electrical engineering (ICITISEE)*, 2017, pp. 294–298. doi: https://doi.org/10.1109/ICITISEE.2017.8285514

[39] G. Bin Huang, Q. Y. Zhu, and C. K. Siew, "EXTREME LEARNING MACHINE: A NEW LEARNING SCHEME OF FEEDFORWARD NEURAL NETWORKS," *IEEE International Conference on Neural Networks - Conference Proceedings*, vol. 2, no. August, pp. 985–990, 2004. doi: https://doi.org/10.1109/IJCNN.2004.1380068

[40] W. Zhang, Z. Zhang, L. Wang, H.-C. Chao, and Z. Zhou, "EXTREME LEARNING MACHINES WITH EXPECTATION KERNELS," *Pattern Recognit*, vol. 96, p. 106960, 2019. doi: https://doi.org/10.1016/j.patcog.2019.07.005

[41] A. L. Afzal, N. K. Nair, and S. Asharaf, "DEEP KERNEL LEARNING IN EXTREME LEARNING MACHINES," *Pattern Analysis and Applications*, vol. 24, pp. 11–19, 2021. doi: https://doi.org/10.1007/s10044-020-00891-8

[42] Google, "GOOGLE COLABORATORY," 2024. [Online]. Available: https://colab.research.google.com