



Weather and Marine Multi-output Prediction Using XGBoost on Automatic Weather Station Data

Willdan Aprizal Arifin^{1*}, Luthfi Anzani², Ma'ruf³, Anton Daud⁴, Lukman Handyanto⁵, Raisa Maulidia⁶, Ramzan Pradana Maulsyid⁷, Angga Fadzar⁸

^{1,2,5,6,7,8}Departement of Marine Information System, Universitas Pendidikan Indonesia, Bandung, Indonesia

³Departement of Maritime Logistics, Universitas Pendidikan Indonesia, Bandung, Indonesia

⁴Meteorology, Climatology, and Geophysics Agency (BMKG) Maritime Meteorological Station Class 1 Merak, Serang, Indonesia

*willdanarifin@upi.edu

Abstract

Climate change on a global scale has triggered an increase in sea levels and heightened the frequency of extreme weather events, especially in maritime countries such as Indonesia. These conditions necessitate the development of accurate and adaptive weather and marine prediction systems. This study proposes a multi-output prediction model using the eXtreme Gradient Boosting (XGBoost) algorithm based on BMKG's Automatic Weather Station (AWS) data from the BMKG. The data cover the period 2022-2025 with high temporal resolution and include five main parameters: wind speed, water level, water temperature, relative humidity, and wind direction. The hyperparameter tuning process led to the discovery of an optimal configuration capable of enhancing the model's accuracy. The evaluation results of the coefficient of determination (R^2) and Root Mean Squared Error (RMSE) metrics show that the model can predict water temperature, water level, and relative humidity with very high accuracy, which is more than 85 percent. The model also performed well in predicting wind speed, although it still faced difficulties in handling wind direction due to its cyclical nature. Overall, the XGBoost approach proved effective in modeling weather and marine parameters simultaneously and has the potential to be integrated into environmental monitoring systems in Indonesia's coastal and archipelagic regions.

Keywords: marine; multioutput prediction; weather; XGBoost

How to Cite: W. A. Arifin, "Weather and Marine Multi-output Prediction Using XGBoost on Automatic Weather Station Data", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 10, no. 2, pp. 257 - 266, Apr. 2026.
Permalink/DOI: <https://doi.org/10.29207/resti.v10i2.7031>

Received: August 1, 2025

Accepted: March 3, 2026

Available Online: April 6, 2026

This is an open-access article under the [CC BY 4.0 License](#)

Published by [Ikatan Ahli Informatika Indonesia](#)

1. Introduction

Climate change is becoming an increasingly urgent global challenge due to its impacts on various aspects of human life around the world [1]. One of the most significant impacts of climate change is sea level rise [2]. As a maritime country, Indonesia faces great challenges in monitoring and predicting dynamic and often changing weather and marine conditions [3]. Information on weather and marine parameters such as wind speed, wind direction, relative humidity, water temperature, and water level is essential to support various sectors, including shipping, fisheries, agriculture, and disaster mitigation. Accurate predictions of these parameters can improve safety and efficiency in decision-making in coastal areas and the high seas [4].

Automatic observation technology, such as Automatic Weather Station (AWS), has been utilized by Meteorology, Climatology, and Geophysics Agency of Indonesia (BMKG) to record meteorological and marine data automatically, in real-time, and continuously [5]. AWS produces data in high temporal resolution that has great potential to be analyzed using machine learning-based predictive approaches [6]. However, the utilization of AWS data still faces a number of challenges, especially related to the complexity of relationships between variables, large data volumes, and interrelationships between parameters that are not always linear. Therefore, algorithms that can handle multivariate data and simultaneously predict multiple outputs are needed.

One relevant algorithm that has been proven effective in dealing with these challenges is eXtreme Gradient Boosting (XGBoost). XGBoost is a tree-based boosting algorithm built with a focus on performance optimization, flexible implementation, and portability. It is capable of producing the best solutions in various data science applications [7]. This algorithm is effective in dealing with complex non-linear relationships and is robust to multicollinearity problems between variables [8]. Research conducted by [9] showed that the XGBoost algorithm effectively handled large-scale data. Considering these advantages, this research develops a multi-output prediction model, which predicts more than one parameter simultaneously, based on historical data from AWS. Furthermore, previous studies in meteorological forecasting show that XGBoost often achieves better prediction performance than several alternative algorithms [10]. These findings provide additional justification for using XGBoost in this study, especially given its suitability for handling complex, multivariate, and cyclical environmental data.

The goal of this research is to build a multi-output predictive model for five key weather and marine parameters, including wind speed, wind direction, relative humidity, water temperature, and water level. Modeling is done using the XGBoost algorithm, which is expected to produce prediction outputs that are more adaptive and responsive to atmospheric and marine dynamics. The results of this research are expected to support early warning systems, data-based decision making, and the development of a more comprehensive weather and marine monitoring system in Indonesia.

2. Methods

The overall stages of this research are summarized in the form of a flowchart, as shown in Figure 1.

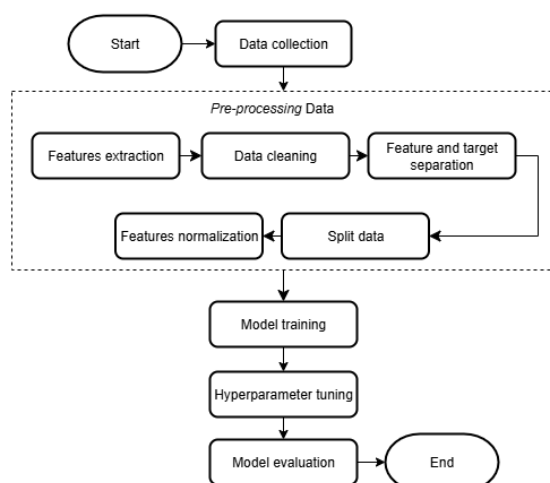


Figure 1. Research flow

As illustrated in Figure 1, this study was conducted through nine systematic stages. The process begins with data collection from Automatic Weather Stations (AWS) as the primary source of weather and marine parameters. The raw data then undergoes

preprocessing, including feature extraction, data cleaning, and separation of features and target variables, to ensure consistency and reliability of the inputs. The dataset is subsequently divided into training and testing subsets, followed by feature normalization to eliminate scale differences across variables. Afterward, the prepared data is used for model training with the XGBoost algorithm, and hyperparameter tuning is performed to optimize its configuration. Finally, the model is evaluated using performance metrics to assess accuracy and generalization capability.

2.1 Data Collection

This research uses data obtained from the AWS tool owned by the BMKG, which functions to record meteorological and marine parameters automatically. The data generated by the AWS system is stored periodically, forming a dataset of environmental observations that can be used for further analysis of atmospheric and oceanic conditions in the study area.

The parameters used in this research include wind speed, wind direction, water temperature, water level, and relative humidity. All data were collected during the period from 2022 to 2025, with a recording interval of one hour, resulting in a high temporal resolution. enabling a more detailed depiction of changes in environmental conditions. This high temporal resolution supports the conduct of more detailed and accurate predictive analyses.

2.2 Features Extraction

Processing time series data and extracting relevant features are widely regarded as two of the most critical stages in traditional machine learning workflows [11]. The feature extraction method is an approach that aims to obtain a new representation of the features in the initial dataset. This technique is very useful for reducing computational burden and simplifying data structures, without losing important information related to prediction [12]. Through feature extraction, existing attributes are processed into new features that are more meaningful, informative, and relevant for the modeling process [13].

In this research, feature extraction is focused on utilizing temporal information from the timestamp variable. Time has cyclical and seasonal characteristics that are important to preserve in weather and marine data modeling. Therefore, time elements such as year, month, date, and hour are extracted as separate features. Furthermore, to capture periodic patterns, the month, day, and hour attributes are converted into sinusoidal and cosinusoidal representations using trigonometric functions. This transformation aims to preserve the cyclic relationship between times, for example between December and January, which are numerically distant (12 and 1) but actually chronologically close.

In addition, additional calendar features such as 'weekday' (day of the week), 'is_weekend' (weekend indicator), and 'season' (season) based on month values

were extracted. Such information is important because atmospheric and marine conditions can fluctuate depending on the time of day, day of the week, or season. These extracted features are then included as input variables in XGBoost modeling, improving the model's ability to understand the time context of changing weather and marine conditions.

2.3 Data Cleaning

Data cleaning is one of the crucial stages in data pre-processing that focuses on identifying and rectifying inaccurate or inconsistent data to improve the quality of the dataset [14]. This stage involves identifying and handling data such as empty or null values [15]. In this research, the data cleaning process focuses on removing entries that have missing values in one or more parameters from both the AWS data and the target data. By ensuring only complete and consistent data is used, this step helps maintain the integrity of the analysis and minimizes the potential for bias or error in the later stages of modeling.

In addition to removing incomplete observations, the cleaning procedure also involved examining temporal continuity and eliminating sensor anomalies that may arise due to calibration drift, communication lag, or environmental interference. AWS data sets are sometimes prone to noise and short-term spikes that do not reflect actual meteorological dynamics [16]. Therefore, ensuring data consistency prior to modeling is essential to prevent bias amplification during training. This step enhances model reliability and contributes to more stable generalization performance.

2.4 Feature and Target Separation

After the cleaning and feature extraction process is completed, the data is separated into two main components, namely input (features) and output (targets). The input features or variables (X) consist of a number of parameters that play a role in describing environmental conditions and time, including geographical coordinates (lon, lat), cyclically transformed time features ('hour_sin', 'hour_cos', 'day_sin', 'day_cos', 'month_sin', 'month_cos'), and additional temporal attributes such as 'weekday', 'is_weekend', and 'season'.

The target variable (y) includes five weather and marine parameters to be predicted simultaneously through a multi-output approach, namely windspeed, water level, water temperature (watertemp), relative humidity (rh), and winddir. This separation aims to make the data have a structure that suits the needs of machine learning model training, especially for the XGBoost algorithm used in this study.

2.5 Split Data

The dataset is divided into training and testing sets during the data splitting process [17], [18]. The purpose of data splitting is to fairly assess the model's performance on data that was not used during training. In this study, data division is done with a proportion of

90% as training data and 10% as test data, utilizing the 'train_test_split' function from the scikit-learn library with a specified random state to ensure consistent and reproducible data splitting results. Training data is used in the model training process so that the algorithm can learn patterns from the relationship between features and targets. Meanwhile, test data is employed to assess how well the trained model is able to accurately predict new data.

The 90:10 data split configuration was selected to maximize the learning capacity of the model while preserving sufficient unseen data for evaluation [19]. Given the relatively large temporal dataset spanning multiple years, this proportion ensures that seasonal and inter-annual variability are adequately represented in both subsets. Such configuration reduces the risk of data fragmentation while maintaining statistical validity in performance assessment.

2.6 Features Normalization

Feature normalization is an important step in data pre-processing, also known as scaling or standardization. This process is intended so that all features are in a balanced range of values [20]. Normalization guarantees that the data maintains a consistent structure and can be accurately processed by machine learning algorithms [21]. In the context of this research, normalization is performed to align the scale of each numeric feature that has a range of different values. This scale imbalance can cause features with small values to be less influential in the learning process than other features that have a larger range, which can reduce model performance [22].

The Standard Scaler method from the scikit learn library is applied to normalize numerical features before model training. Using Standard Scaler involves converting the values of each feature into a scale that reflects their relative distance from the mean, measured in terms of standard deviations [23]. The Standard Scaler is used to convert numerical features in the dataset thus ensuring they are centered at zero with a unit standard deviation, ensuring a more uniform and balanced data distribution [24]. The main purpose of this process is to align the distribution scale between features, so that each feature has a balanced weight and can be processed more efficiently by machine learning algorithms [25]. This process begins by calculating statistics on the training data, which is then used to transform both training and test data. Normalization is only used for the input features 'X_train' and 'X_test', while the target variable (y) remains in its original scale. The purpose of this step is to ensure that each feature has an equal contribution to the model training, accelerate convergence, and improve stability and prediction accuracy.

2.7 Model Training

XGBoost was introduced by [26], a machine learning method that constructs an ensemble of decision trees sequentially. XGBoost utilizes a decision tree algorithm

as the core model and integrates boosting techniques to gradually and accurately improve prediction performance [27]. The model training process in this study was conducted using the XGBoost algorithm in a multi-output regression framework. Mathematically, the XGBoost algorithm is formulated as shown in Equation 1 [28], [29].

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (1)$$

$\hat{y}_i$ represents the forecasted value corresponding to the i -th data point. In this equation, K represents the total number of decision trees used in the model. The term $f_k(x_i)$ refers to the function of the k -th regression tree that generates an output based on the input features. Each f_k belongs to the set F , which contains all possible regression tree functions that can be formed within the model.

The model is trained to predict multiple weather and marine parameters at once, namely windspeed, wind direction, relative humidity (rh), water temperature (watertemp), and water level. To handle multiple outputs, a 'MultiOutputRegressor' approach is used, which wraps the XGBoost model and allows separate regression training for each target.

The normalized training data is used for the model fitting process. This process involves training the input feature data (extracted time and location features) and the target data. The core parameters of the model, namely the number of trees (`n_estimators`), maximum depth (`max_depth`), and learning rate (`learning_rate`), are adjusted during the training process based on the tuning results to form the best-performing model. Training is performed on the training data ('`X_train_scaled`' and '`y_train`') that has gone through a thorough pre-processing stage.

2.8 Hyperparameter Tuning

The hyperparameter tuning method deals with the process of exploring possible combinations of hyperparameter values in the model architecture, with the aim of finding the optimal configuration that produces the best prediction performance [30]. The application of this tuning makes the model more robust in the face of data variations and more stable in producing consistent predictions [31]. Unlike model parameters, which are derived from the training data, hyperparameters must be specified beforehand and carefully tuned to ensure optimal model performance [32]. In this study, the tuning process is applied to the XGBoost model packaged in the Multi Output Regressor approach, allowing simultaneous modeling of several weather and marine target variables. Some of the important hyperparameters adjusted include the number of decision trees ('`n_estimators`'), the maximum depth of the tree ('`max_depth`'), the learning rate ('`learning_rate`'), as well as the L1 ('`reg_alpha`') and L2 ('`reg_lambda`') regularization parameters. The adjustment of these values is crucial as it directly affects the accuracy and quality of the predictions produced by the model [33].

The tuning process is performed using Grid Search and Randomized Search approaches with a 3-fold CV cross-validation technique. GridSearchCV is one of the features in the scikit learn library that allows the hyperparameter search process to be performed automatically and systematically [34]. For the multi-output model, tuning is done by inserting estimator parameters in each XGBoost hyperparameter so that it can be processed in the context of the MultiOutputRegressor wrapper. In addition, a custom score function based on the average coefficient of determination (R^2) is used, so that the model is not only optimal at one target, but is able to provide balanced performance across all target variables. The tuning process produces the best model based on the highest validation score, which is then used in the evaluation stage.

2.9 Model Evaluation

The evaluation stage is carried out to assess the extent to which the performance of the XGBoost model that has been developed is able to accurately predict weather and marine parameters [35]. This evaluation serves to measure the generalization ability of the model to the test data, as well as ensuring that the model not only works well on the training data, but can also be relied upon to predict new data consistently. The evaluation process includes two main approaches, namely quantitative evaluation and visual evaluation. In quantitative evaluation, Root Mean Squared Error (RMSE) and the coefficient of determination (R^2) are used. RMSE is calculated as formulated in Equation 2 [36].

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

RMSE measures the average magnitude of prediction errors in the same units as the target variable, where y_i is the observed value for each data point, $\hat{y}_i$ is the predicted value, and n is the total number of data points. The coefficient of determination (R^2) is calculated using Equation 3 [36].

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (3)$$

R^2 indicates the proportion of variance in the observed data that is explained by the model, where y_i is the observed value, $\hat{y}_i$ is the predicted value, and $\bar{y}$ is the mean of all observed values. Values closer to 1 indicate better predictive performance. Additionally, the difference between the R^2 values of the training and test datasets (overfitting gap) was calculated to detect potential overfitting.

Visual evaluation is done by plotting a comparison graph between the actual and predicted values for each target parameter. This graph provides an intuitive picture of the model's consistency and precision in following the data pattern. From the evaluation results, the model performed well with high R^2 values and low RMSE for most of the target parameters, indicating that

the model is capable of performing multi-output predictions with sufficient accuracy.

3. Results and Discussions

3.1 Hyperparameter Tuning Results

Model evaluation was carried out to appraise the predictive performance of the optimized XGBoost model on five target parameters: windspeed, water level, watertemp, relative humidity (RH), and wind direction. The evaluation uses the R^2 metric to assess the prediction accuracy as well as the RMSE as an indicator of the average prediction error. The results of the evaluation are presented according to the performance on training and testing data. The best parameter combinations obtained from the tuning process are presented in Table 1.

Table 1. Results of hyperparameter tuning

Parameters	Grid Search Value	Best Value
'n_estimators'	1000, 1200, 1500	1500
'max_depth'	8, 9, 10	10
'learning_rate'	0.03, 0.05, 0.07	0.05
'subsample'	0.9, 0.95, 1.0	0.9
'colsample_bytree'	0.85, 0.9, 0.95	0.95
'reg_alpha'	0.5, 0.7, 1.0	0.5
'reg_lambda'	0.5, 1.0, 1.5	0.5

A high 'n_estimators' parameter indicates that the model requires many decision trees to achieve stable and accurate prediction results. According to [37], the higher the 'n_estimators' value, the more boosting iterations are performed to obtain optimal performance. Meanwhile, a large maximum tree depth ('max_depth')

value indicates that the model learns complex patterns in the data. The deeper the tree structure, the more complex the patterns learned by the model [38]. To avoid overfitting due to such complexity, regularization techniques such as 'reg_alpha' and 'reg_lambda' are used to control the model weights [39].

In addition, using a learning rate that is not too high allows the model to learn slowly and steadily. This is important because the learning rate governs how much adjustment is made to the model's weights and parameters during the training process [40]. The 'subsample' and 'colsample_bytree' parameters, which are both close to 1, indicate that almost all training data and features are used in each learning iteration, while still leaving room for variation between trees to improve model generalization. The combination of these parameters creates a model that is balanced between complexity and adaptability to the data and is ready to be used in the training process of multi-output models.

3.2 Model Evaluation Results

The XGBoost model was applied to predict five weather and marine variables: windspeed, wind direction, water level, water temperature, and relative humidity. The evaluation is done using the RMSE metric and the R^2 , and is presented in the form of a graph comparing the actual value with the predicted value for each variable. Overall, the model produced R^2 values ranging from 0.71 to 0.99, indicating very high accuracy for some parameters while showing more challenges for others. Figure 2 specifically illustrates the evaluation results for the windspeed variable.

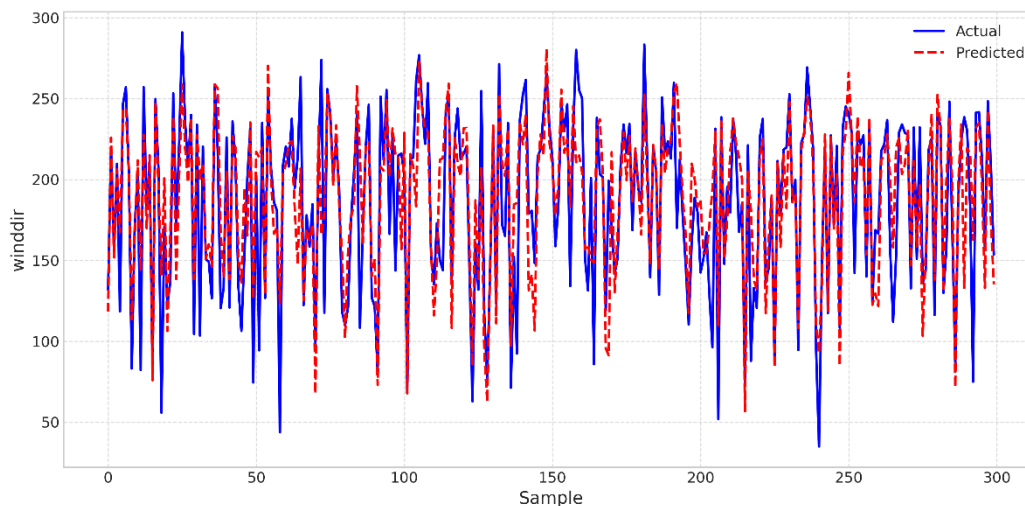


Figure 2. Graph of evaluation results on the windspeed variable

For the wind speed variable, the model produced an RMSE value of 0.81 and R^2 of 0.82, highlighting that the model was capable of describing about 82% of the variation in the wind speed data. As seen in Figure 2, the prediction line (red) generally follows the pattern of the actual line (blue), although there are small deviations, especially at the peaks of extreme values. This shows that the model is quite sensitive to seasonal

changes or short-term fluctuations in wind speed, although its accuracy decreases slightly at very high values. Compared to other parameters, windspeed shows a relatively balanced performance, not as stable as water level but significantly better than wind direction. Meanwhile, the evaluation results for the water level variable are shown in Figure 3.

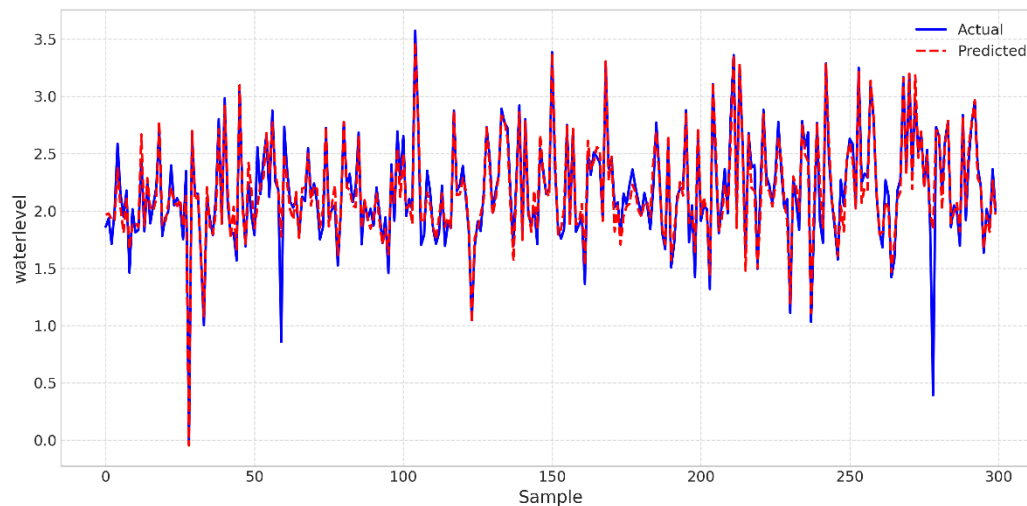


Figure 3. Graph of evaluation results on the water level variable

The predictions for the water level variable performed very well, with an RMSE of 0.13 and an R^2 of 0.93. Figure 3 shows that the prediction line almost completely merges with the actual line, indicating a high agreement between the prediction and observation data. This indicates that the XGBoost model is very effective in studying sea level rise and fall patterns, which tend to be more stable than other atmospheric variables. The performance here is noticeably stronger compared to windspeed, confirming that XGBoost is particularly suited for variables with consistent temporal behavior.

The model shows very accurate prediction results on the water temperature variable, with an RMSE of 0.59 and a very high R^2 value of 0.99. Figure 4, the red line (prediction) almost completely follows the blue line (actual) with very little deviation. This performance can be attributed to the relatively more stable seawater temperature characteristics that are easily studied by the model due to consistent seasonal or daily patterns. Among all variables, water temperature is the most predictable, showing that the XGBoost model can capture highly structured and stable patterns with almost perfect accuracy.

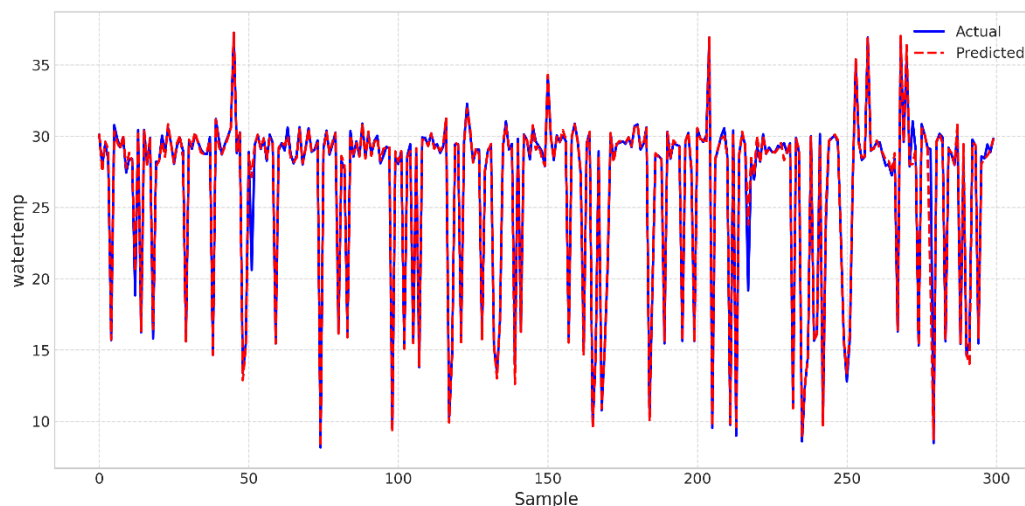


Figure 4. Graph of evaluation results on the watertemp variable

The evaluation of relative humidity shows an RMSE of 2.63 and an R^2 of 0.85, indicating that the model does a good job of modeling this variable. The fourth graph in Figure 5 shows that the prediction line follows the direction of change of the actual value quite well, although there are small deviations at extreme points. The fluctuating humidity values may make it difficult for the model to capture the entire pattern with perfect precision, but the overall accuracy remains high. This result also places relative humidity performance

between water temperature and windspeed, confirming that XGBoost performs best when variability is moderate but not excessive.

Wind direction is the lowest performing variable in this model, with an RMSE of 29.54 and an R^2 of only 0.71. From the final graph in Figure 6, it can be seen that the prediction line tends to deviate more significantly than other variables, especially in the case of sudden changes in wind direction.

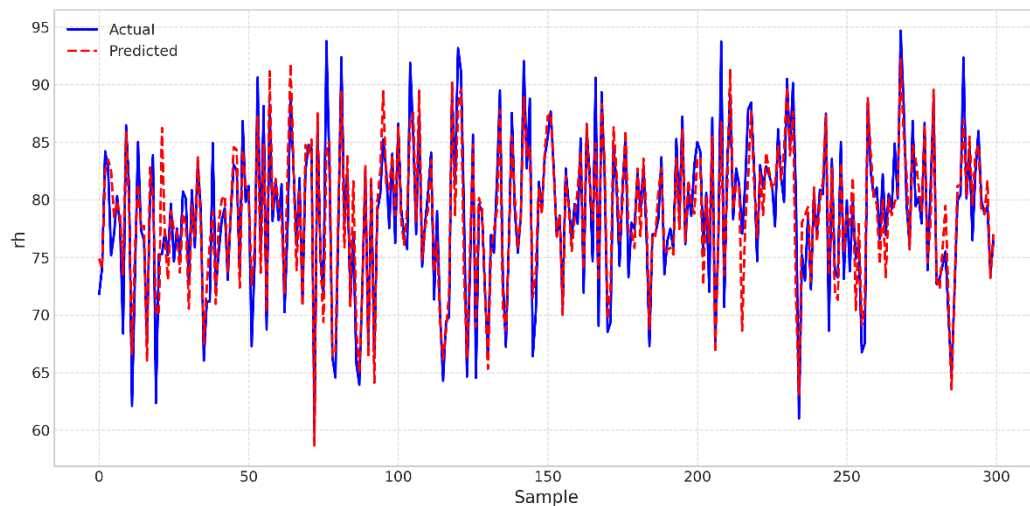


Figure 5. Graph of evaluation results on the rh variable

Overall, the XGBoost model showed very satisfactory prediction performance, especially for the watertemp, water level and relative humidity (rh) variables, which each obtained an R^2 value above 0.85. This result indicates that the model is able to capture the patterns and relationships in the data well for these three parameters. Furthermore, the difference between training and testing R^2 values was monitored to evaluate potential overfitting. The relatively small performance

gap indicates that the model maintains consistent predictive capability across unseen data [41]. This suggests that the applied regularization parameters effectively controlled model complexity despite the high number of boosting iterations. Such stability is essential for operational deployment in real-time maritime monitoring systems.

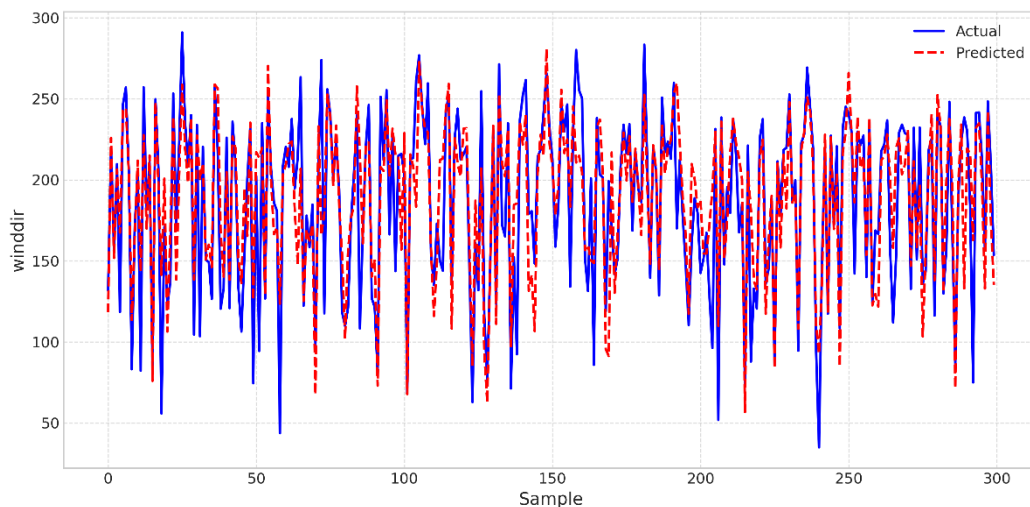


Figure 6. Graph of evaluation results on the winddir variable

However, the model performance tends to decrease when applied to variables that are characteristically more dynamic and complex, such as wind direction. The lower R^2 value for this parameter reflects the challenge of modeling wind direction behavior, which is highly influenced by high spatial and temporal variability, as well as dependence on external factors that are not fully captured in the dataset. This is similar to the study by [42], which found similar limitations in wind-based forecasts. The study showed that although XGBoost was able to compete with Random Forest and DAN2, this model was consistently outperformed by CNN in predicting wind speeds under various weather conditions in South Africa, confirming the difficulty of

modeling wind-based variables. The results of [43] also show that XGBoost consistently outperforms Random Forest and classical regression in rainfall forecasting, with R^2 values above 0.80. These findings are consistent with the results of this study on water temperature and water level variables (R^2 0.85–0.89), which show that XGBoost performs best on variables with a relatively stable temporal structure. Furthermore, [10] confirm the superiority of XGBoost over Random Forest and Support Vector Regression (SVR) in predicting weather parameters in Indonesia, with an R^2 achievement of 0.8183. However, the study also highlights a decline in performance when predicting wind direction, a finding that is in line with the results of this study.

The comparison results not only reinforce the methodological justification for using XGBoost but also explain the specific conditions under which this model excels or is less than optimal. Beyond methodological comparisons, this study also has practical implications for coastal and maritime operations. The high accuracy obtained for water temperature and sea level highlights the potential application of XGBoost in supporting operational services such as early warning systems, maritime safety management, and port operations. Reliable forecasts for these parameters can help reduce the risks faced by small-scale fishermen and coastal communities, while providing useful decision support for government agencies and stakeholders involved in maritime transportation and disaster risk management.

Furthermore, this study contributes to the development of machine learning applications for weather and maritime forecasting by applying the XGBoost algorithm in a multi-output framework using AWS data. Unlike most previous studies that focused on single-output predictions or conventional statistical approaches, this study demonstrates XGBoost's ability to capture multiple interdependent target variables simultaneously. The results also provide empirical evidence of the algorithm's robustness to relatively stable parameters, while revealing limitations in cyclical variables such as wind direction. Therefore, this study highlights the need for further research to explore hybrid approaches, such as combining boosting algorithms with sequential deep learning models or cyclic transformations, as well as incorporating additional data sources such as satellite datasets and reanalysis to improve the generalization and accuracy of predictions.

3.3 Performance Under Extreme Maritime Conditions

The reliability of the model in predicting uncertain maritime phenomena was evaluated through testing under extreme conditions. These extreme conditions were determined using a statistical distribution approach (data-driven approach) due to the limited availability of explicit labels for extreme weather

events. A percentile-based approach was used to define extreme events as values that exceed a certain distribution threshold and are rare [44]. In this study, the 95th percentile threshold was used to define extreme values for the variables of wind speed, water level, and water temperature. The relative humidity variable involves a range of values below the 5th percentile and above the 95th percentile to capture air humidity anomalies. Table 2 presents the results of the RMSE and R^2 metric evaluations specifically for the dataset under these extreme conditions.

The model shows stable performance on the waterlevel variable with an R^2 value of 0.96 and an RMSE of 0.12, indicating the model's ability to maintain prediction accuracy under extreme water level conditions. The watertemp prediction produces the highest R^2 of 0.99, reflecting the ability of the XGBoost algorithm to capture ocean thermal patterns even at peak temperature ranges. The relative humidity variable also shows good performance ($R^2 = 0.92$), indicating the model's reliability in representing coastal humidity dynamics. For the windspeed variable, an RMSE of 0.85 with an R^2 of 0.88 was obtained, indicating an increase in error compared to the evaluation of the entire dataset, along with increased variability in strong wind conditions. The largest prediction error was found in the wind direction variable (winddir) with an RMSE of 28.22 because wind direction is a circular variable that inherently causes numerical discontinuities in conventional regression approaches [45].

Table 2. Error Metrics under Extreme Conditions

Variables	RMSE	R^2
Windspeed	0.85	0.88
Waterlevel	0.12	0.96
Watertemp	0.72	0.99
Relative Humidity	2.89	0.92
Wind Direction	28.22	0.79

To provide a visual illustration of the model's ability to follow actual data fluctuations across all output variables under extreme conditions, a time-series comparison is presented in Figure 7.

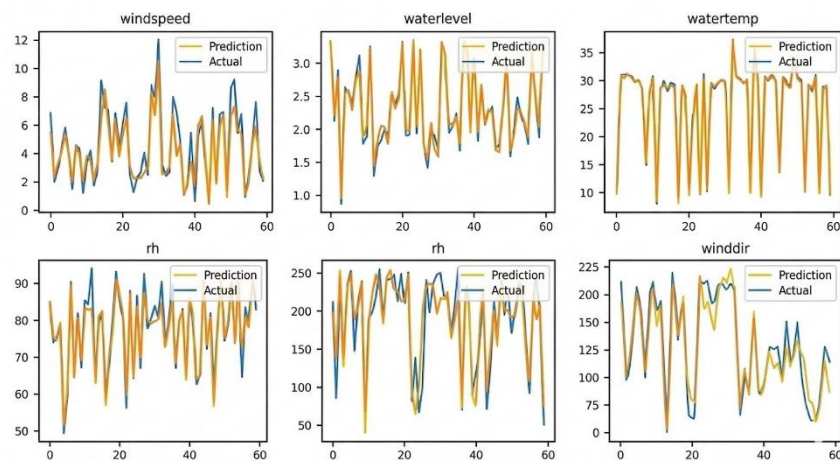


Figure 7. Comparison of fluctuations in predicted vs actual data under extreme conditions

The visualization in Figure 7 shows that the model is able to follow the dynamics of multi-output variables under extreme conditions, with prediction lines that are generally consistent with the observation patterns across all variables. The highest conformity is seen in watertemp and waterlevel, in line with R^2 values approaching 1, while relative humidity and windspeed continue to show adaptive capabilities despite an increase in deviation during sharp fluctuations. The wind direction variable shows wider residuals due to its circular and volatile nature. Overall, despite an increase in error in several variables, the consistently high R^2 value (>0.79) indicates that the gradient boosting approach has good robustness in dealing with high variability in extreme maritime conditions.

4. Conclusions

This research successfully developed a multi-output prediction model based on the XGBoost algorithm to estimate five weather and marine parameters at once, namely windspeed, water level, watertemp, relative humidity, and wind direction. Systematic hyperparameter adjustments resulted in the most suitable combination of parameters to optimize the model. The model showed excellent prediction performance, especially for the watertemp, water level, and RH variables, with R^2 values above 0.85 and low prediction error rates.

Model evaluation shows that XGBoost is very effective in learning patterns of relatively stable variables, such as water temperature and water level. However, the model faces challenges in predicting wind direction. Nonetheless, the prediction results for all variables remain within an acceptable level of accuracy for practical applications. In general, the multi-output XGBoost approach proved to be reliable in handling the simultaneous prediction of various weather and marine parameters, and has the potential to be integrated into AWS data-based environmental monitoring systems in Indonesia's coastal and archipelagic regions.

References

- [1] G. E. Hayatulah, J. Mahasari, M. Ihsan, M. B. A. Wicaksono, and S. Alhamda, "Kebijakan Lingkungan dalam Menanggapi Permasalahan Perubahan Iklim di Indonesia: Sebuah Tinjauan Integratif," *Jurnal Birokrasi & Pemerintahan Daerah*, vol. 5, no. 2, pp. 266–276, 2023, doi: [10.15575/jbpd.v5i2.28548](https://doi.org/10.15575/jbpd.v5i2.28548)
- [2] B. D. Dasanto, S. Sulistiyanti, A. Anria, and R. Boer, "Dampak Perubahan Iklim terhadap Kenaikan Muka Air Laut di Wilayah Pesisir Pangandaran," *Risalah Kebijakan Pertanian dan Lingkungan*, vol. 7, no. 2, pp. 82–94, 2020, doi: [10.29244/jkebijakan.v7i2.28039](https://doi.org/10.29244/jkebijakan.v7i2.28039)
- [3] J. Risandi et al., "Hydrodynamics Across Seagrass Meadows and Its Impacts on Indonesian Coastal Ecosystems: A Review," 2023, *Frontiers Media SA*, doi: [10.3389/feart.2023.1034827](https://doi.org/10.3389/feart.2023.1034827)
- [4] A. Durap, "Interpretable Machine Learning for Coastal Wind Prediction: Integrating SHAP Analysis and Seasonal Trends," *J. Coast. Conserv.*, vol. 29, 2025, doi: [10.1007/s11852-025-01108-y](https://doi.org/10.1007/s11852-025-01108-y)
- [5] N. M. A. Suandewi, A. Muliantara, and M. A. Raharja, "Pengembangan Sistem Database Lokasi Automatic Rain Gauge (ARG) dan Automatic Weather Stations (AWS) Berbasis Website di Balai Besar MKG Wilayah III Denpasar," *Jurnal Pengabdian Informatika*, vol. 1, no. 4, pp. 1153–1160, 2023, , [Online]. Available: <https://ejournal1.unud.ac.id/index.php/jupita/article/view/423>
- [6] G. Fragiadakis, E. Filiopoulou, C. Michalakelis, T. Kamalakis, and M. Nikolaidou, "Applying Machine Learning in Cloud Service Price Prediction: The Case of Amazon IaaS," *Future Internet*, vol. 15, no. 8, 2023, doi: [10.3390/fi15080277](https://doi.org/10.3390/fi15080277)
- [7] S. Hakkal and A. A. Lahcen, "XGBoost To Enhance Learner Performance Prediction," *Computers and Education: Artificial Intelligence*, vol. 7, 2024, doi: [10.1016/j.caeai.2024.100254](https://doi.org/10.1016/j.caeai.2024.100254)
- [8] A. Rahman, V. Paramarta, A. N. Ida, M. Harits Akbar, V. Sonya, and M. Simanjuntak, "Perbandingan Kinerja SVR dan XGBoost untuk Peramalan Emisi CO₂ Global Berbasis Machine Learning," *Jurnal Komtika (Komputasi dan Informatika)*, vol. 9, no. 1, pp. 37–44, 2025, doi: [10.31603/komtika.v9i1.13449](https://doi.org/10.31603/komtika.v9i1.13449)
- [9] P. S. Rizky, R. H. Hirzi, and U. Hidayaturohman, "Perbandingan Metode LightGBM dan XGBoost dalam Menangani Data dengan Kelas Tidak Seimbang," *J Statistika*, vol. 15, no. 2, pp. 228–236, 2022, [Online]. Available: www.unipasby.ac.id
- [10] A. Syahreza, N. K. Ningrum, and M. A. Syahrasy, "Perbandingan Kinerja Model Prediksi Cuaca: Random Forest, Support Vector Regression, dan XGBoost," *Edumatic: Jurnal Pendidikan Informatika*, vol. 8, no. 2, pp. 526–534, 2024, doi: [10.29408/edumatic.v8i2.27640](https://doi.org/10.29408/edumatic.v8i2.27640)
- [11] J. Van Der Donckt, J. Van Der Donckt, E. Deprost, and S. Van Hoecke, "tsflex: Flexible time series processing & feature extraction," *SoftwareX*, vol. 17, 2022, doi: [10.1016/j.softx.2021.100971](https://doi.org/10.1016/j.softx.2021.100971)
- [12] H. Faiqoh, "Penerapan Ensemble Feature Selection untuk Mengurangi Dimensi dalam Prediksi Data Time Series," Skripsi, Universitas Islam Negeri Syarif Hidayatullah Jakarta, Tangerang Selatan, 2023.
- [13] D. Rohadi, "Pengaruh Ensemble Feature Selection pada Prediksi Data Time Series Menggunakan Gated Recurrent Unit (GRU) dan Bidirectional Long Short-Term Memory (Bi-LSTM) (Studi Kasus: Walmart)," Skripsi, Universitas Islam Negeri Syarif Hidayatullah Jakarta, Tangerang Selatan, 2024.
- [14] A. M. Sharifnia, D. E. Kpormegbey, D. K. Thapa, and M. Cleary, "A Primer of Data Cleaning in Quantitative Research: Handling Missing Values and Outliers," *J. Adv. Nurs.*, 2025, doi: [10.1111/jan.16908](https://doi.org/10.1111/jan.16908)
- [15] E. S. Nugraha, Z. Alike, and D. A. Hamzah, "Forecasting the Stock Price of PT Astra International Using the LSTM Method," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 8, no. 3, pp. 431–437, Jun. 2024, doi: [10.29207/resti.v8i3.5699](https://doi.org/10.29207/resti.v8i3.5699)
- [16] I. D. G. A. Putra et al., "Development of Typical Meteorological Years Based on Quality Control of Datasets in Indonesia," *Renew. Energy*, vol. 221, 2024, doi: [10.1016/j.renene.2023.119699](https://doi.org/10.1016/j.renene.2023.119699)
- [17] A. Yasin, A. Yuniarti, and Y. Adi Nugroho, "Efektifitas Algoritma Data Mining dalam Menentukan Pendorong Darah Potensial," *Syntax J. Inf.*, vol. 11, no. 01, pp. 12–22, Jun. 2022, doi: [10.35706/syji.v11i01.6595](https://doi.org/10.35706/syji.v11i01.6595)
- [18] R. Oktafiani, A. Hermawan, and D. Avianto, "Pengaruh Komposisi Split Data Terhadap Performa Klasifikasi Penyakit Kanker Payudara Menggunakan Algoritma Machine Learning," *Jurnal Sains dan Informatika*, vol. 9, no. 1, pp. 19–28, 2023, doi: [10.34128/jsi.v9i1.622](https://doi.org/10.34128/jsi.v9i1.622)
- [19] Q. H. Nguyen et al., "Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil," *Math. Probl. Eng.*, 2021, doi: [10.1155/2021/4832864](https://doi.org/10.1155/2021/4832864)
- [20] A. Demircioğlu, "The Effect of Feature Normalization Methods in Radiomics," *Insights Imaging*, vol. 15, no. 2, pp. 1–11, Dec. 2024, doi: [10.1186/s13244-023-01575-7](https://doi.org/10.1186/s13244-023-01575-7)
- [21] H. Karmila and S. Yuliyatini, "Systematic Literature Review: Metode Normalisasi Data yang Efektif pada Deteksi Penyakit Diabetes Menggunakan Machine Learning (Algoritma K-Nearest Neighbors)," *Journal of Sciencetech Research and Development*, vol. 6, no. 1, pp. 2099–2110, 2024, doi: [10.56670/jsrd.v6i1.436](https://doi.org/10.56670/jsrd.v6i1.436)
- [22] P. P. Alloserung, A. Erna, M. Bagussahrir, and S. Alam, "Analisis Performa Normalisasi Data untuk Klasifikasi K-Nearest Neighbor pada Dataset Penyakit," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 9, no. 3, pp. 178–191, 2024, doi: [10.14421/jiska.2024.9.3.178-191](https://doi.org/10.14421/jiska.2024.9.3.178-191)

- [23] Z. Wahyuzi, A. Luthfi, and D. H. Fudholi, "Predicting Smart Office Electricity Consumption in Response to Weather Conditions Using Deep Learning," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 8, no. 1, pp. 100–110, Feb. 2024, doi: [10.29207/resti.v8i1.5530](https://doi.org/10.29207/resti.v8i1.5530).
- [24] M. Saputra, J. P. Sidabuke, R. P. Sinulingga, and R. B. Tamba, "Analisis Metode Algoritma K-Nearest Neighbor (KNN) dan Naive Bayes untuk Klasifikasi Diabetes Mellitus," *Jurnal TEKNIKOM*, vol. 6, no. 2, pp. 723–729, 2023, doi: [10.37600/tekinkom.v6i2.942](https://doi.org/10.37600/tekinkom.v6i2.942).
- [25] M. F. R. Mahendra, N. L. Azizah, and S. Sumarno, "Implementasi Machine Learning untuk Memprediksi Cuaca Menggunakan Support Vector Machine," *Jurnal Ilmiah KOMPUTASI*, vol. 23, no. 1, pp. 45–50, 2024, doi: [10.32409/jikstik.23.1.3449](https://doi.org/10.32409/jikstik.23.1.3449).
- [26] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Aug. 2016, pp. 785–794, doi: [10.1145/2939672.2939785](https://doi.org/10.1145/2939672.2939785).
- [27] M. A. Rayadin, M. Musaruddin, R. A. Saputra, and I. Isnawaty, "Implementasi Ensemble Learning Metode XGBoost dan Random Forest untuk Prediksi Waktu Penggantian Baterai Aki," *BIOS: Jurnal Teknologi Informasi dan Rekayasa Komputer*, vol. 5, no. 2, pp. 111–119, 2024, doi: [10.37148/bios.v5i2.128](https://doi.org/10.37148/bios.v5i2.128).
- [28] K. D. K. Wardani and M. Akbar, "Diabetes Risk Prediction using Feature Importance Extreme Gradient Boosting (XGBoost)," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 4, pp. 824–831, 2023, doi: [10.29207/resti.v7i4.4651](https://doi.org/10.29207/resti.v7i4.4651).
- [29] R. Winurputra and D. E. Ratnawati, "Peramalan Penjualan Produk Menggunakan Extreme Gradient Boosting (XGBoost) dan Kerangka Kerja CRISP-DM untuk Pengoptimalan Manajemen Persediaan (Studi Kasus: UB Mart)," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 2, pp. 417–428, 2025, doi: [10.25126/jtiik.2025129451](https://doi.org/10.25126/jtiik.2025129451).
- [30] T. A. E. Putri, T. Widiarini, and R. Santoso, "Penerapan Tuning Hyperparameter RandomSearchCV pada Adaptive Boosting untuk Prediksi Kelangsungan Hidup Pasien Gagal Jantung," *Jurnal Gaussian*, vol. 11, no. 3, pp. 397–406, 2023, doi: [10.14710/j.gauss.11.3.397-406](https://doi.org/10.14710/j.gauss.11.3.397-406).
- [31] A. Amalia, M. Radhi, D. R. H. Sitompul, S. H. Sinurat, and E. Indra, "Prediksi Harga Mobil Menggunakan Algoritma Regresi dengan Hyperparameter Tuning," *JUSIKOM PRIMA (Jurnal Sistem Informasi dan Ilmu Komputer Prima)*, vol. 4, no. 2, pp. 28–32, 2021, doi: [10.34012/jurnalsisteminformasidanilmukomputer.v4i2.2479](https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v4i2.2479).
- [32] J. A. Ilemobayo et al., "Hyperparameter Tuning in Machine Learning: A Comprehensive Review," *Journal of Engineering Research and Reports*, vol. 26, no. 6, pp. 388–395, 2024, doi: [10.9734/jerr/2024/v26i61188](https://doi.org/10.9734/jerr/2024/v26i61188).
- [33] H. Mulyo and A. K. Zyen, "Pengaruh Hyperparameter Tuning Gradient Boosting Terhadap Prediksi Pemilihan Program Studi Mahasiswa Baru," *Bulletin of Computer Science Research*, vol. 5, no. 2, pp. 131–137, 2025, doi: [10.47065/bulletincsr.v5i2.454](https://doi.org/10.47065/bulletincsr.v5i2.454).
- [34] Z. M. E. Darmawan and A. F. Dianta, "Implementasi Optimasi Hyperparameter GridSearchCV Pada Sistem Prediksi Serangan Jantung Menggunakan SVM," *Teknologi: Jurnal Ilmiah Sistem Informasi*, vol. 13, no. 1, pp. 8–15, 2023, doi: [10.26594/teknologi.v13i1.3098](https://doi.org/10.26594/teknologi.v13i1.3098).
- [35] M. F. R. Aditya, N. L. Azizah, and U. Indahyanti, "Prediksi Penyakit Hipertensi Menggunakan Metode Decision Tree dan Random Forest," *Jurnal Ilmiah Komputasi*, vol. 23, no. 1, pp. 9–16, 2024, doi: [10.32409/jikstik.23.1.3503](https://doi.org/10.32409/jikstik.23.1.3503).
- [36] 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, pp. 1–24, 2021, doi: [10.7717/PEERJ-CS.623](https://doi.org/10.7717/PEERJ-CS.623).
- [37] G. Chairunisa et al., "Life Expectancy Prediction Using Decision Tree, Random Forest, Gradient Boosting, and XGBoost Regressions," *Jurnal Sintak*, vol. 2, no. 2, pp. 71–82, 2024, doi: [10.62375/jsintak.v2i2.249](https://doi.org/10.62375/jsintak.v2i2.249).
- [38] E. J. Sudarman and S. Budi, "Pengembangan Model Kecerdasan Mesin Extreme Gradient Boosting untuk Prediksi Keberhasilan Studi Mahasiswa," *Jurnal Strategi*, vol. 5, no. 2, pp. 297–314, 2023, [Online]. Available: <https://mail.strategi.it.maranatha.edu/index.php/strategi/article/view/437>.
- [39] M. B. Ilmi and K. Kusriani, "Perbandingan Kinerja Algoritma Machine Learning dalam Deteksi Potensi Risiko HIV," *Jurnal Buffer Informatika*, vol. 11, no. 1, pp. 36–46, 2025, [Online]. Available: <https://journal.fkom.uniku.ac.id/buffer>.
- [40] S. Suryadi and M. Masrizal, "XGBoost Algorithm for Cervical Cancer Risk Prediction: Multi-dimensional Feature Analysis," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 9, no. 3, pp. 535–541, 2025, doi: [10.29207/resti.v9i3.6587](https://doi.org/10.29207/resti.v9i3.6587).
- [41] S. Nguyen, X. Fu, D. Ogawa, and Q. Zheng, "An Application-Oriented Testing Regime and Multi-Ship Predictive Modeling for Vessel Fuel Consumption Prediction," *Transp. Res. E Logist. Transp. Rev.*, vol. 177, 2023, doi: [10.1016/j.tre.2023.103261](https://doi.org/10.1016/j.tre.2023.103261).
- [42] F. W. Mugware, C. Sigauke, and T. Ravele, "Evaluating Wind Speed Forecasting Models: A Comparative Study of CNN, DAN2, Random Forest and XGBOOST in Diverse South African Weather Conditions," *Forecasting*, vol. 6, no. 3, pp. 672–699, 2024, doi: [10.3390/forecast6030035](https://doi.org/10.3390/forecast6030035).
- [43] V. Kumar, N. Kedam, O. Kisi, S. Alsulamy, K. M. Khedher, and M. A. Salem, "A Comparative Study of Machine Learning Models for Daily and Weekly Rainfall Forecasting," *Water Resources Management*, vol. 39, no. 1, pp. 271–290, 2025, doi: [10.1007/s11269-024-03969-8](https://doi.org/10.1007/s11269-024-03969-8).
- [44] L. Brunner and A. Voigt, "Pitfalls in Diagnosing Temperature Extremes," *Nat. Commun.*, vol. 15, 2024, doi: [10.1038/s41467-024-46349-x](https://doi.org/10.1038/s41467-024-46349-x).
- [45] H. Shu, W. Song, Y. Wang, and J. Zhang, "WaveHiTS: Wavelet-Enhanced Hierarchical Time Series Modeling for Wind Direction Nowcasting in Eastern Inner Mongolia," *ArXiv*, 2025, doi: [10.48550/arXiv.2504.06532](https://doi.org/10.48550/arXiv.2504.06532).