JTIF : Journal of Technology Information
Volume 02 No. 03, November 2025 p 18 – 28
E-ISSN: 3036-0088

INTERPRETABLE CARBON NANO SENSOR MODELING FOR WATER POLLUTION CLASSIFICATION

Junervin [*1], Jaenudin[2], Syamsuwarni Rambe[3], Silmi Azmi[4], Muhammad Luqmanul Hakim[5], Amina Kurniasi Alu[6]

[1,3,4,5,6] Agroindustrial Engineering, Faculty of Computer Science and Engineering, Universitas Linggabuana PGRI Sukabumi, Indonesia
[2]Informatics, Faculty of Computer Science and Engineering, Universitas Linggabuana PGRI Sukabumi, Indonesia

Email: [1]junervin@unlip.ac.id, [2]jaenudin.fkip.pti@gmail.com, [3]s.rambe@unlip.ac.id, [4]silmiazmi@unlip.ac.id, [5]mluqmanulhakim@unlip.ac.id, [6]amina.alu@unlip.ac.id

Article Info

Article history:
Received November, 02, 2025
Revised November, 19, 2025
Accepted November, 31, 2025

Keywords:
Decision support,
electrochemical analytics,
explainable artificial intelligence,
predictive modeling,
signal interpretation.

Abstract

Electrochemical carbon nano sensors are promising for rapid water pollution monitoring, but translating multivariate sensor responses into reliable pollution level decisions remains challenging. This study develops an interpretable machine learning framework for electrochemical carbon nano sensor response modeling in water pollution classification. A public environmental nanosensor dataset containing 6,786 observations was used, with graphene, multi walled carbon nanotube based, and hybrid carbon nanomaterial sensor records. The target variable consisted of two pollution severity classes, namely Low and Medium. Random Forest, Extreme Gradient Boosting, Support Vector Machine, and Multilayer Perceptron were evaluated using accuracy, precision, recall, F1 score, confusion matrix, receiver operating characteristic analysis, and precision recall analysis. The Multilayer Perceptron achieved the highest accuracy of 98.38 percent and macro F1 score of 96.62 percent, while Extreme Gradient Boosting achieved 97.86 percent accuracy and was selected for deployment. Interpretability analysis identified lead concentration, mercury concentration, nitrogen dioxide concentration, benzene concentration, rolling mean, and rolling standard deviation as dominant predictors. The proposed framework supports explainable water pollution classification and provides a deployable prototype for intelligent environmental monitoring.

Corresponding Author:
Junervin
Agroindustrial Engineering, Faculty of Computer Science and Engineering, Universitas Linggabuana PGRI Sukabumi, Indonesia
Email: junervin@unlip.ac.id

1. INTRODUCTION

Water pollution remains one of the most persistent environmental problems because aquatic systems receive contaminants from domestic, agricultural, and industrial activities. Conventional laboratory based analysis can provide accurate measurements, but it often requires expensive instrumentation, trained personnel, complex sample preparation, and delayed reporting. These limitations reduce its suitability for continuous and decentralized monitoring, particularly when rapid warning is required to support environmental management. In this context, sensor based monitoring has become an important direction because it enables faster observation of water quality indicators and supports more responsive decision making.

Electrochemical sensors have attracted considerable attention for water quality monitoring due to their sensitivity, portability, relatively low cost, simple operation, and compatibility with on site analysis. Recent studies have shown that electrochemical sensing platforms can be developed for a wide range of water contaminants, including heavy metals, nutrients, phenolic compounds, pharmaceutical residues, and other emerging pollutants [1], [2]. The performance of these sensors is strongly influenced by electrode surface properties, charge transfer behavior, analyte interaction, and environmental conditions such as pH, dissolved oxygen, and water conductivity. Therefore, the interpretation of electrochemical signals should not rely only on single response values, but should consider multivariate sensor response

patterns. Recent studies further show that electrochemical sensing platforms have been increasingly used for contaminants of emerging concern, heavy metal ions, lead detection, multiplexed sensing, and machine learning integrated water monitoring systems [7]–[13]. These developments indicate that the integration of electrochemical sensing and computational intelligence is becoming an important direction in intelligent water quality monitoring.

Carbon nanomaterials are particularly important in electrochemical sensing because of their high surface area, electrical conductivity, chemical stability, and favorable electron transfer properties. Graphene, carbon nanotubes, and hybrid carbon nanomaterials have been widely explored as electrode modifiers for water contaminant detection [3], [4]. Their structural and electronic characteristics can improve signal amplification and enhance the sensitivity of the sensing interface. However, the same advantages also introduce modeling complexity. Sensor responses generated from carbon nanomaterial based platforms may vary according to material type, functionalization, surface area, conductivity, pollutant class, and temporal fluctuations. This complexity creates a need for computational models that can learn nonlinear relationships between sensor attributes, electrochemical signals, environmental parameters, pollutant concentrations, and pollution severity levels. Recent reviews also emphasize the role of nanomaterial based electrochemical detection and biosensing platforms in improving environmental quality monitoring [5], [6].

Machine learning offers a promising approach for addressing this challenge. Instead of depending exclusively on fixed threshold rules, machine learning models can extract patterns from high dimensional sensor data and transform them into predictive decisions. Artificial intelligence has increasingly been applied in environmental monitoring to improve pollution detection, risk prediction, and water quality assessment [14]–[19]. In sensor analytics, supervised learning algorithms can be used to classify pollution levels, estimate contaminant concentrations, and identify important variables associated with environmental risk. Nevertheless, predictive accuracy alone is not sufficient for scientific and environmental applications. A model that provides high accuracy but limited interpretability may be difficult to trust, particularly when the output is intended to support environmental monitoring or policy related decisions. Recent water quality studies have demonstrated the effectiveness of supervised machine learning, ensemble learning, and deep learning for predicting water quality indicators, assessing uncertainty, and identifying influential physicochemical variables [14]–[25]. These studies support the use of data driven models for environmental monitoring, but most focus on conventional water quality parameters rather than electrochemical carbon nano sensor response modeling.

Interpretability is therefore an essential requirement in machine learning based sensor response modeling. Interpretable models and post hoc explanation methods help clarify which variables contribute most to a prediction. In the context of electrochemical carbon nano sensors, interpretability can reveal whether pollutant concentration features, electrochemical peak features, material characteristics, or temporal statistics dominate classification outcomes. This is important because the scientific value of the model does not only depend on whether it predicts the correct pollution class, but also on whether its decision pattern is consistent with plausible sensor behavior and environmental mechanisms. Feature importance and permutation importance can support this objective by quantifying the contribution of individual variables to model performance and prediction behavior [26].

Although many studies have reviewed electrochemical sensors for water pollutant detection, fewer studies have focused on the complete data driven workflow that connects carbon nanomaterial sensor response modeling, interpretable machine learning, comparative model evaluation, and deployable decision support. Existing studies often emphasize sensor fabrication, analytical performance, or contaminant detection capability, while less attention is given to how multivariate response data can be transformed into a reproducible classification model and implemented as an accessible web based prototype. This gap is important because practical environmental monitoring systems require not only accurate sensing materials, but also computational models that can process sensor responses, explain classification decisions, and deliver results in a usable interface.

This study addresses that gap by developing an interpretable machine learning framework for electrochemical carbon nano sensor response modeling in water pollution level classification. A public Electrochemical Carbon Nano Sensor Dataset available from Kaggle was used as the experimental basis [28]. The dataset contains sensor material characteristics, electrochemical response features, environmental water parameters, pollutant concentration values, temporal response statistics, and categorical pollution level labels. The target variable consists of two pollution severity classes, namely Low and Medium. Several machine learning models were evaluated, including Random Forest, Extreme Gradient Boosting, Support Vector Machine, and Multilayer Perceptron. Model performance was assessed using accuracy, precision, recall, F1 score, confusion matrix, receiver operating characteristic analysis, and precision recall analysis. Interpretability was examined through feature importance and permutation importance, with the most practical high performing model deployed through a Streamlit based web application.

The main contribution of this study is threefold. First, it presents a reproducible machine learning workflow for classifying water pollution levels from electrochemical carbon nano sensor response data. Second, it provides interpretability analysis to identify the dominant predictors associated with pollution level classification. Third, it demonstrates a deployable web based prototype that translates the trained model into an accessible decision support tool for sensor based environmental monitoring. By integrating predictive modeling, interpretability, and deployment, this study contributes to the development of intelligent water quality monitoring systems that are not only accurate but also explainable and practically usable.

2. METHOD

This study applied a quantitative experimental approach based on supervised machine learning to model electrochemical carbon nano sensor responses for water pollution level classification. The workflow consisted of dataset acquisition, data inspection, preprocessing, feature transformation, model training, model evaluation, interpretability analysis, and web based deployment. The overall objective was to transform multivariate sensor response data into a

predictive and explainable classification model that can distinguish water pollution severity levels.

2.1 Dataset and Research Variables

The experimental data were obtained from a public Electrochemical Carbon Nano Sensor Dataset available on Kaggle, containing 6,786 observations and 27 columns [28]. The dataset represents electrochemical sensing records generated from carbon nanomaterial based sensing platforms for water pollution monitoring. The sensor material variables included material type, surface area, electrical conductivity, and surface functionalization. The electrochemical response variables included peak current, peak voltage, charge transfer resistance, capacitance, cyclic voltammetry area, differential pulse voltammetry peak height, and signal to noise ratio. Environmental water variables included temperature, pH, dissolved oxygen, and water conductivity. Pollutant concentration variables consisted of lead concentration, mercury concentration, nitrogen dioxide concentration, and benzene concentration. The dataset also included pollutant class and temporal response variables, namely previous reading, rolling mean with a window of five readings, and rolling standard deviation with a window of five readings.

The target variable was Pollution Level. Although the dataset description refers to pollution severity categories, the actual data used in this study contained two target classes, namely Low and Medium. Therefore, the task was formulated as a binary classification problem. The Low class consisted of 5,841 samples, while the Medium class consisted of 945 samples. This distribution indicates class imbalance, which was considered during model evaluation by reporting macro averaged precision, recall, and F1 score in addition to overall accuracy.

Table 1. Feature Groups and Research Variables.

Feature Group	Variables	Role in Modeling
Sensor material characteristics	Material_Type, Surface_Area, Conductivity, Functionalization	Represents carbon nanomaterial properties that may influence sensor behavior
Electrochemical response features	Peak_Current, Peak_Voltage, Charge_Transfer_Resistance, Capacitance, CV_Area, DPV_Peak_Height, SNR	Represents electrochemical signal response patterns generated during measurement
Environmental water parameters	Temperature, pH, Dissolved_Oxygen, Water_Conductivity	Represents water condition during sensor measurement
Pollutant concentration variables	Pb_Concentration, Hg_Concentration, NO2_Concentration, Benzene_Concentration	Represents contaminant intensity associated with pollution severity
Temporal response statistics	Previous_Reading, Rolling_Mean_5, Rolling_Std_5	Represents short term response dynamics and recent pollution trends
Categorical pollutant information	Pollutant_Class	Represents the broad pollutant category in the sample
Target variable	Pollution_Level	Represents the pollution severity class to be predicted

2.2 Data Preprocessing

The preprocessing stage was designed to ensure that the dataset could be used consistently across machine learning algorithms. Three non predictive columns were removed from the modeling input, namely Sensor ID, Timestamp, and Time Index. Sensor ID was excluded because it only represented the identity of each reading. Timestamp and Time Index were excluded from the main classification model to avoid direct dependency on record order and to keep the experiment focused on sensor response attributes and pollutant related variables.

The remaining variables were divided into numerical and categorical features. The numerical features consisted of sensor material measurements, electrochemical response values, environmental water parameters, pollutant concentrations, and temporal response statistics. The categorical features consisted of Material Type, Functionalization, and Pollutant Class. Categorical variables were transformed using one hot encoding to preserve category level information without imposing ordinal assumptions. After encoding, the number of input features increased to 28. The target label was transformed using label encoding, where the two target classes were converted into numerical class codes for model training.

Feature scaling was performed using standardization. Each numerical and encoded feature was transformed to have zero mean and unit variance. Standardization was applied because several algorithms, particularly Support Vector Machine and Multilayer Perceptron, are sensitive to feature scale. The dataset was then divided into training and testing subsets using an 80 to 20 split. Stratified splitting was applied to preserve the proportion of the Low and Medium classes in both subsets. The final training set contained 5,428 samples, while the testing set contained 1,358 samples.

2.3 Machine Learning Models

Four supervised learning algorithms were evaluated to compare ensemble, boosted tree, margin based, and neural network based approaches, namely Random Forest [29], Extreme Gradient Boosting [30], Support Vector Machine [31], and Multilayer Perceptron trained through back propagation [32]. Random Forest was used as an ensemble tree based baseline because it can model nonlinear relationships and handle heterogeneous feature types after encoding. Extreme Gradient Boosting was used as a gradient boosted decision tree model because of its strong predictive capability, regularization strategy, and practical suitability for tabular data. Support Vector Machine with a radial basis function kernel was used to evaluate margin based nonlinear classification. Multilayer Perceptron was used as a neural network model to learn nonlinear interactions among sensor response variables.

Random Forest was configured with 300 estimators and balanced class weighting. Extreme Gradient Boosting was configured for binary logistic classification with 300 estimators, a learning rate of 0.05, maximum tree depth of 5,

subsampling ratio of 0.8, and column sampling ratio of 0.8. Support Vector Machine used a radial basis function kernel, probability estimation, and balanced class weighting. The Multilayer Perceptron architecture consisted of dense layers with 128, 64, and 32 neurons, batch normalization, dropout regularization, and a softmax output layer with two neurons. Early stopping was applied by monitoring validation loss to reduce overfitting.

Table 2. Model Configuration.

Model	Main Configuration	Preprocessing Requirement	Purpose in Experiment
Random Forest	300 estimators, balanced class weighting, random state 42	One hot encoding and feature scaling	Tree based ensemble baseline for nonlinear classification
Extreme Gradient Boosting	300 estimators, learning rate 0.05, maximum depth 5, subsample 0.8, column sample by tree 0.8, binary logistic objective	One hot encoding and feature scaling	High performing boosted tree model for tabular data and deployment
Support Vector Machine	Radial basis function kernel, probability estimation enabled, balanced class weighting, random state 42	One hot encoding and feature scaling	Margin based nonlinear classifier for comparison
Multilayer Perceptron	Dense layers with 128, 64, and 32 neurons, batch normalization, dropout regularization, softmax output, early stopping	One hot encoding and feature scaling	Neural network model for learning nonlinear feature interactions

2.4 Model Evaluation

The models were evaluated on the independent testing set. The evaluation metrics included accuracy, macro precision, macro recall, macro F1 score, weighted precision, weighted recall, and weighted F1 score. Accuracy was used to measure the overall proportion of correctly classified samples. However, because the target distribution was imbalanced, macro averaged metrics were emphasized to assess whether the models performed well across both classes. Macro F1 score was particularly important because it balances precision and recall equally across classes regardless of class frequency.

Confusion matrix analysis was used to inspect correct and incorrect classifications for each class. Receiver operating characteristic analysis was used to evaluate the discriminatory capability of the models in binary classification [33]. Precision recall analysis was also conducted because it is informative when class imbalance exists [34]. The model comparison was used to identify the best performing model for predictive performance and the most practical model for deployment.

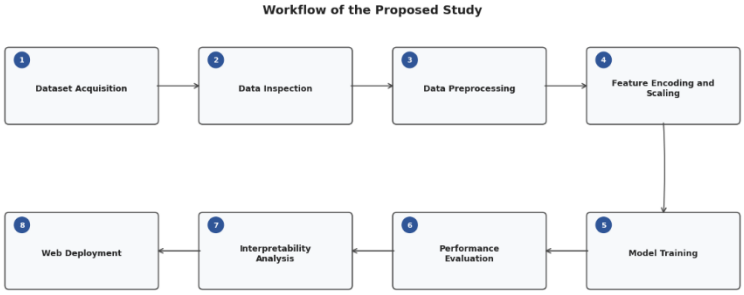

Figure 1. Workflow of the proposed study, including dataset acquisition, data inspection, preprocessing, feature transformation, model training, performance evaluation, interpretability analysis, and Streamlit based deployment.

2.5 Interpretability Analysis

Interpretability analysis was conducted to identify variables that contributed most to pollution level classification. Two complementary methods were used. First, model based feature importance was extracted from tree based models to identify features that contributed to split decisions during training. Second, permutation importance was applied to the selected model by measuring the decrease in macro F1 score after randomly permuting each feature [26]. This method provides a model agnostic estimate of feature relevance and is useful for validating whether the most important variables are consistent with the sensor response and pollutant concentration patterns. Explainable artificial intelligence methods are increasingly used to improve trust in machine learning predictions by providing local or global explanations of model behavior [35]. Tree based explanation approaches such as SHAP are also relevant for future model interpretation, although the present study reports model based feature importance and permutation importance as the main interpretation outputs [27].

The interpretability analysis focused on identifying whether electrochemical response variables, pollutant concentration variables, environmental water parameters, or temporal response statistics were the dominant predictors. This step was important because the objective of the study was not only to build an accurate classifier, but also to understand which sensor related and pollutant related variables were most influential in the prediction process.

2.6 Web Based Deployment

After model comparison, the high performing and deployment friendly model was serialized together with the feature scaler, label encoder, feature name list, and metadata. These components were integrated into a Streamlit web application available at https://carbon-sensor-app.streamlit.app/. The web prototype allows users to input sensor material characteristics, electrochemical response values, environmental parameters, pollutant concentrations, pollutant class, and temporal statistics. The application processes the input using the same preprocessing pipeline used during training and returns the predicted pollution level with class probability values.

The deployment was included to demonstrate the practical usability of the proposed framework as a decision support prototype. The web application does not replace laboratory based water quality testing. Instead, it demonstrates how trained machine learning models can be integrated into an accessible interface for rapid simulation, demonstration, and future

development of intelligent water pollution monitoring systems.

3. RESULTS AND DISCUSSION

This section presents the experimental results obtained from the electrochemical carbon nano sensor dataset, including exploratory data analysis, class distribution, model performance comparison, interpretability analysis, and web based implementation. The results are discussed in relation to the objective of developing an interpretable machine learning framework for water pollution level classification from multivariate sensor response data.

3.1 Dataset Characteristics and Class Distribution

The dataset contained 6,786 sensor observations and 27 original columns. After removing non predictive identifiers and temporal index variables, the modeling process used sensor material characteristics, electrochemical response variables, environmental water parameters, pollutant concentration values, pollutant class, and temporal response statistics. The dataset included 20 numerical features and 3 categorical features after excluding the target variable and non predictive columns. The categorical variables consisted of Material Type, Functionalization, and Pollutant Class. After one hot encoding, the final model input consisted of 28 features.

The target variable was Pollution Level. The actual dataset used in this study contained two pollution severity classes, namely Low and Medium. The Low class consisted of 5,841 samples, while the Medium class consisted of 945 samples. This indicates a clear class imbalance, with the Low class representing approximately 86.07 percent of the dataset and the Medium class representing approximately 13.93 percent. This class distribution has important implications for model evaluation. Accuracy alone may provide an overly optimistic view because a model can achieve high accuracy by favoring the majority class. Therefore, macro precision, macro recall, and macro F1 score were included to evaluate whether the model could classify both pollution levels more fairly.

Table 3. Dataset Characteristics.

Dataset Attribute	Value
Number of observations	6786
Number of original columns	27
Number of removed non predictive columns	3
Removed columns	Sensor_ID, Timestamp, Time_Index
Number of numerical input features	20
Number of categorical input features	3
Number of encoded input features	28
Number of missing values	2222
Number of duplicated rows	0
Target variable	Pollution_Level
Target classes	Low, Medium
Class distribution	Low: 5841 samples (86.07%); Medium: 945 samples (13.93%)

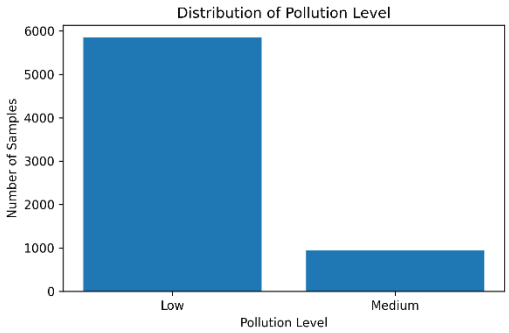

Figure 2. Distribution of pollution level classes, showing the dominance of the Low class compared with the Medium class.

The material distribution was relatively balanced across the three carbon nanomaterial categories. Graphene appeared in 2,303 observations, Hybrid material appeared in 2,250 observations, and multi walled carbon nanotubes appeared in 2,233 observations. This relatively even distribution suggests that the dataset provides comparable representation of different carbon nanomaterial based sensing platforms. The pollutant class distribution was also balanced, with Gas appearing in 2,276 observations, Organic pollutants in 2,261 observations, and Heavy Metal pollutants in 2,249 observations. This distribution supports the use of the dataset for evaluating sensor response modeling across diverse pollutant categories.

Cross tabulation between Material Type and Pollution Level showed that the Low class dominated across all material categories. Graphene contained 1,972 Low samples and 331 Medium samples. Hybrid materials contained 1,940 Low samples and 310 Medium samples. MWCNT contained 1,929 Low samples and 304 Medium samples. A similar pattern appeared for pollutant classes. Gas contained 1,946 Low samples and 330 Medium samples, Heavy Metal contained 1,963 Low samples and 286 Medium samples, and Organic contained 1,932 Low samples and 329 Medium samples. These results indicate that class imbalance was not limited to a single material type or pollutant class, but was present across the dataset.

3.2 Exploratory Feature Analysis

Exploratory analysis was conducted to examine the relationship between numerical variables and the encoded pollution level. The strongest correlation with the target was observed for lead concentration, with a correlation coefficient of 0.5258. This result suggests that lead concentration was the most linearly associated variable with the transition from Low to Medium pollution level. Rolling mean with a window of five readings showed the second strongest association with the target, with a correlation coefficient of 0.2953. Mercury concentration, rolling standard deviation, nitrogen dioxide concentration, and benzene concentration also showed relevant associations, with correlation coefficients of 0.2414, 0.1913, 0.1869, and 0.1474 respectively.

These findings show that the target class was more strongly associated with pollutant concentration variables and temporal response statistics than with material identity alone. This is scientifically plausible because pollution level classification should be influenced by contaminant intensity and recent sensor response behavior. The importance of rolling mean and rolling standard deviation also indicates that temporal context contributes to the classification process. Even though the main model did not use Timestamp and Time Index, temporal statistical features preserved useful information regarding previous sensor response trends.

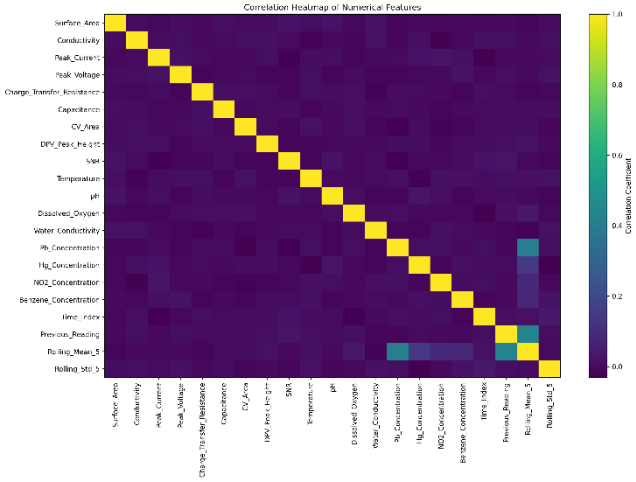

Figure 3. Correlation heatmap of numerical features used to examine relationships among sensor responses, pollutant concentrations, environmental parameters, and temporal statistics.

Table 4. Top Feature Correlations with Pollution Level.

Feature	Correlation with Encoded Pollution Level
Pb_Concentration	0.5258
Rolling_Mean_5	0.2953
Hg_Concentration	0.2414
Rolling_Std_5	0.1913
NO2_Concentration	0.1869
Benzene_Concentration	0.1474

Dimensionality reduction using principal component analysis was used to visualize the distribution of Low and Medium classes in two dimensional space. The PCA visualization showed partial separation between the two classes, although some overlap remained. This indicates that the pollution level classification task is not purely linearly separable in low dimensional space. Therefore, nonlinear machine learning algorithms are suitable for this problem because they can model complex interactions among pollutant concentrations, sensor responses, and environmental variables.

3.3 Classification Performance

Four machine learning models were evaluated on the testing set, namely Random Forest, Extreme Gradient Boosting, Support Vector Machine, and Multilayer Perceptron. The Multilayer Perceptron achieved the best overall performance, with an accuracy of 0.9838 and a macro F1 score of 0.9662. Extreme Gradient Boosting achieved the second highest performance, with an accuracy of 0.9786 and a macro F1 score of 0.9541. Support Vector Machine obtained an accuracy of 0.9669 and a macro F1 score of 0.9363. Random Forest produced the lowest result among the evaluated models, with an accuracy of 0.9514 and a macro F1 score of 0.8856.

Table 5. Classification Performance Comparison.

Model	Accuracy	Macro Precision	Macro Recall	Macro F1 Score	Weighted Precision	Weighted Recall	Weighted F1 Score
Multilayer Perceptron	0.9838	0.9662	0.9662	0.9662	0.9838	0.9838	0.9838
Extreme Gradient Boosting	0.9786	0.9684	0.9410	0.9541	0.9784	0.9786	0.9783
Support Vector Machine	0.9669	0.9066	0.9741	0.9363	0.9721	0.9669	0.9681
Random Forest	0.9514	0.9474	0.8431	0.8856	0.9510	0.9514	0.9482

Figure 4. Comparison of model performance based on accuracy, macro precision, macro recall, and macro F1 score.

The superior performance of the Multilayer Perceptron indicates that nonlinear neural network based learning was effective for capturing the relationships between multivariate sensor variables and pollution level categories. The use of dense layers, batch normalization, dropout, and early stopping likely contributed to the stable learning process. However, Extreme Gradient Boosting also achieved a very competitive result and provided practical advantages for deployment. It is well suited for tabular data, supports probability estimation, provides feature importance values, and can be serialized efficiently for integration with web based applications. Therefore, Extreme Gradient Boosting was selected for the deployed Streamlit prototype, even though the Multilayer Perceptron produced the highest numerical accuracy.

The use of macro F1 score is important in interpreting these results. Because the dataset was imbalanced, high accuracy alone would not sufficiently indicate whether the minority Medium class was identified correctly. The high macro F1 scores obtained by the Multilayer Perceptron and Extreme Gradient Boosting indicate that both models were able to classify the two classes with strong balance between precision and recall. This is important for water pollution monitoring because failure to detect elevated pollution levels can reduce the usefulness of the system as an early decision support tool.

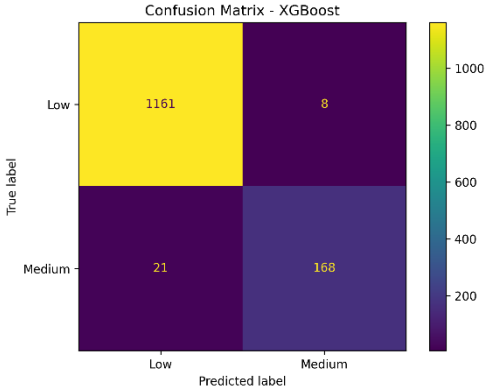

Figure 5. Confusion matrix of the Extreme Gradient Boosting model on the independent testing set.

Receiver operating characteristic analysis and precision recall analysis were also used to evaluate the discriminatory capability of the models. The ROC curve comparison should demonstrate the ability of each model to distinguish between Low and Medium pollution classes across classification thresholds. The precision recall curve is particularly relevant because the Medium class is less frequent than the Low class. A strong precision recall curve indicates that the model can maintain reliable positive class detection under class imbalance.

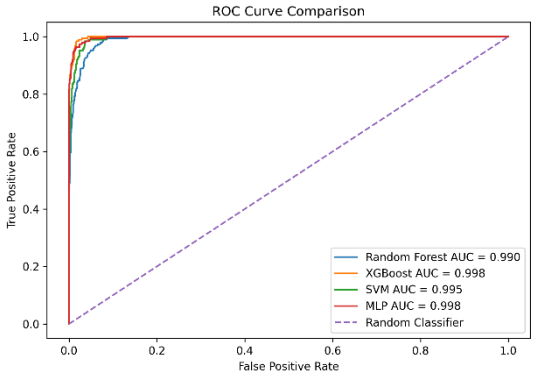

Figure 6. Receiver operating characteristic curve comparison of Random Forest, Extreme Gradient Boosting, Support Vector Machine, and Multilayer Perceptron models.

3.4 Interpretability of the Selected Model

Interpretability analysis was conducted to determine which variables contributed most to the classification of pollution level. The feature importance results from Extreme Gradient Boosting showed that pollutant concentration variables were the most influential predictors. Lead concentration produced the highest importance score of 0.3140, followed by mercury concentration with 0.1169, nitrogen dioxide concentration with 0.0883, rolling mean with 0.0819, benzene concentration with 0.0713, and rolling standard deviation with 0.0527.

Permutation importance produced a consistent pattern. Lead concentration was again the most influential feature, with an importance score of 0.3862. Mercury concentration ranked second with 0.1682, followed by nitrogen dioxide concentration with 0.1181 and benzene concentration with 0.0763. These results confirm that pollutant concentration variables had the strongest effect on classification performance. The consistency between model based importance and permutation importance strengthens the validity of the interpretation because two different explanation methods identified similar dominant predictors.

Table 6. Top Influential Features.

Feature	XGBoost Importance	Permutation Importance	Permutation Std	Interpretation
Pb_Concentration	0.3140	0.3862	0.0156	Dominant heavy metal indicator associated with pollution severity
Hg_Concentration	0.1169	0.1682	0.0154	Heavy metal indicator contributing to class separation
NO2_Concentration	0.0883	0.1181	0.0122	Inorganic pollutant indicator associated with pollution intensity
Rolling_Mean_5	0.0819	0.0008	0.0039	Temporal average representing recent pollution response trend
Benzene_Concentration	0.0713	0.0763	0.0086	Organic pollutant indicator contributing to pollution classification
Rolling_Std_5	0.0527	-0.0015	0.0041	Temporal variation representing short term response fluctuation

The dominance of pollutant concentration features is reasonable because the classification target represents pollution severity. Lead concentration emerged as the most important predictor, indicating that variations in lead concentration contributed strongly to the distinction between Low and Medium pollution levels. Mercury, nitrogen dioxide, and benzene concentrations also contributed to class separation, suggesting that the model used multiple contaminant indicators rather than relying on a single pollutant. The contribution of rolling mean and rolling standard deviation further shows that temporal response statistics helped summarize recent pollution related trends. These findings support the idea that interpretable machine learning can provide insight into both contaminant intensity and sensor response dynamics.

Electrochemical response variables such as peak current, peak voltage, charge transfer resistance, capacitance, cyclic voltammetry area, differential pulse voltammetry peak height, and signal to noise ratio remained relevant as part of the overall input representation, although the highest importance scores were concentrated in pollutant concentration and temporal variables. This result suggests that the dataset label structure may be strongly driven by concentration values. For future experimental sensor datasets, it would be valuable to further examine whether raw electrochemical signal features alone can classify pollution severity when direct pollutant concentration values are unavailable.

3.5 Web Based Prototype Deployment

The selected Extreme Gradient Boosting model was deployed as a Streamlit based web application at https://carbon-sensor-app.streamlit.app/. The deployed application allows users to input sensor material properties, electrochemical response values, environmental water parameters, pollutant concentrations, pollutant class, and temporal response statistics. The application then applies the same preprocessing components used during model training, including feature encoding alignment, scaling, label decoding, and class probability estimation. The output includes the predicted pollution level and probability values for each class.

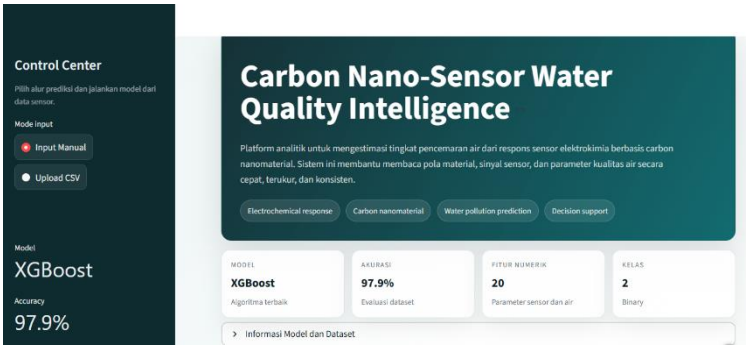

Figure 7. Streamlit dashboard interface for manual input, model information, pollution level prediction, and probability visualization.

The deployment demonstrates that machine learning based sensor response modeling can be translated into a usable decision support interface. This is important because many machine learning studies stop at model evaluation and do not show how the model can be accessed by end users. The web prototype provides a bridge between computational modeling and practical environmental monitoring. Nevertheless, the application should be interpreted as a computational prototype rather than a replacement for certified laboratory analysis or a validated field monitoring system. Since the dataset is public and computational, the system is best viewed as a proof of concept for intelligent sensor data interpretation. Further validation using real experimental measurements from fabricated electrochemical carbon nano sensors would be necessary

before practical field deployment.

4. DISCUSSION

The results indicate that machine learning can effectively classify water pollution levels from electrochemical carbon nano sensor related variables. This finding is consistent with previous studies showing that machine learning and deep learning can improve water quality prediction, uncertainty assessment, and sensor based contaminant detection [11]–[19]. The strong performance of both Multilayer Perceptron and Extreme Gradient Boosting suggests that nonlinear models are suitable for multivariate sensor datasets. Similar trends have been reported in water quality prediction studies where ensemble models, support vector machines, and deep learning models captured nonlinear relationships among physicochemical parameters [20]–[25]. In this study, interpretability analysis further showed that the most influential predictors were chemically and environmentally relevant variables, particularly lead, mercury, nitrogen dioxide, benzene concentration, and temporal response statistics.

The findings imply that combining pollutant concentrations, electrochemical responses, material properties, environmental parameters, and temporal statistics can support data driven classification of pollution severity. Interpretability is also essential because environmental monitoring models should not only provide accurate predictions, but also explain which variables drive their decisions. The Streamlit deployment further demonstrates how a trained model can be translated into an accessible prototype for environmental decision support. However, several limitations remain. The dataset was obtained from a public source and may not fully represent field measurements under uncontrolled conditions. The target variable contained only two classes, Low and Medium, and the dataset was imbalanced. In addition, pollutant concentration variables were included as input features, which improves classification performance but may reduce the extent to which the model represents prediction from raw electrochemical responses alone. Future studies should validate the framework using real experimental sensor data, compare models using all variables versus electrochemical response and environmental features only, expand pollution severity categories, and integrate uncertainty estimation for risk informed monitoring.

5. CONCLUSION

This study developed an interpretable machine learning framework for electrochemical carbon nano sensor response modeling in water pollution classification. The public environmental nanosensor dataset contained 6,786 observations and two classes, namely Low and Medium. The Multilayer Perceptron achieved the highest performance, with 98.38 percent accuracy and 96.62 percent macro F1 score. Extreme Gradient Boosting achieved 97.86 percent accuracy and was selected for Streamlit deployment because of its practical integration capability and interpretability support. Feature importance and permutation importance identified lead concentration, mercury concentration, nitrogen dioxide concentration, benzene concentration, rolling mean, and rolling standard deviation as dominant predictors. These findings indicate that interpretable machine learning can support explainable classification of water pollution levels from carbon nano sensor related data.

DATA AVAILABILITY

The dataset used in this study is publicly available from Kaggle under the Electrochemical Carbon Nano Sensor Dataset [28]. The trained prototype is available through the Streamlit deployment described in the manuscript. The generated code, preprocessing artifacts, and model files can be made available by the corresponding author upon reasonable request.

DECLARATIONS

The authors declare no conflict of interest. This study used a public dataset and did not involve human participants, animal subjects, or private personal data.

REFERENCES

[1] I. Yaroshenko, D. Kirsanov, M. Marjanovic, P. A. Lieberzeit, O. Korostynska, A. Mason, I. Frau, and A. Legin, “Real-time water quality monitoring with chemical sensors,” Sensors, vol. 20, no. 12, Art. no. 3432, 2020, doi: 10.3390/s20123432.

[2] O. Kanoun, T. Lazarević-Pašti, I. Pašti, S. Nasraoui, M. Talbi, A. Brahem, A. Adiraju, E. Sheremet, R. D. Rodriguez, M. Ben Ali, and A. Al-Hamry, “A review of nanocomposite-modified electrochemical sensors for water quality monitoring,” Sensors, vol. 21, no. 12, Art. no. 4131, 2021, doi: 10.3390/s21124131.

[3] G. Cho, S. Azzouzi, G. Zucchi, and B. Lebental, “Electrical and electrochemical sensors based on carbon nanotubes for the monitoring of chemicals in water. A review,” Sensors, vol. 22, no. 1, Art. no. 218, 2022, doi: 10.3390/s22010218.

[4] S. R. Benjamin and E. J. Miranda Ribeiro Júnior, “Graphene-based electrochemical sensors for detection of environmental pollutants,” Current Opinion in Environmental Science and Health, vol. 29, Art. no. 100381, 2022, doi: 10.1016/j.coesh.2022.100381.

[5] H. A. Potes-Lesoinne, F. Ramirez-Alvarez, V. H. Perez-Gonzalez, S. O. Martinez-Chapa, and R. C. Gallo-Villanueva, “Nanomaterials for electrochemical detection of pollutants in water. A review,” Electrophoresis, vol. 43, no. 1 and 2, pp. 249–262, 2022, doi: 10.1002/elps.202100204.

[6] S. Gavrilaș, C. Ș. Ursachi, S. Perța-Crișan, and F. D. Munteanu, “Recent trends in biosensors for environmental quality monitoring,” Sensors, vol. 22, no. 4, Art. no. 1513, 2022, doi: 10.3390/s22041513.

[7] C. Gibi, K. G. S. Kumar, S. K. Sahoo, M. R. Sujith, and A. K. Nayak, "Recent advances on electrochemical sensors for detection of contaminants of emerging concern," Molecules, vol. 28, no. 23, Art. no. 7916, 2023, doi: 10.3390/molecules28237916.
[8] S. F. Sulthana, S. R. Sahoo, N. S. Kumar, A. K. Nayak, and P. K. Sahoo, "Electrochemical sensors for heavy metal ion detection in aqueous medium: A systematic review," ACS Omega, vol. 9, no. 24, pp. 25493–25512, 2024, doi: 10.1021/acsomega.4c00933.
[9] J. Jjagwe, P. W. Olupot, R. Kulabako, and S. Carrara, "Electrochemical sensors modified with iron oxide nanoparticles nanocomposites for voltammetric detection of Pb II in water. A review," Heliyon, vol. 10, no. 8, Art. no. e29743, 2024, doi: 10.1016/j.heliyon.2024.e29743.
[10] M. Ateia, H. Wei, and S. Andreescu, "Sensors for emerging water contaminants: Overcoming roadblocks to innovation," Environmental Science and Technology, vol. 58, no. 6, pp. 2636–2651, 2024, doi: 10.1021/acs.est.3c09889.
[11] S. A. Lahari, N. Kumawat, K. Amreen, R. N. Ponnalagu, and S. Goel, "IoT integrated and deep learning assisted electrochemical sensor for multiplexed heavy metal sensing in water samples," npj Clean Water, vol. 8, Art. no. 10, 2025, doi: 10.1038/s41545-025-00441-x.
[12] B. S. Alfurhood, M. S. M. Danthuluri, S. Jadhav, B. Mouleswararao, N. V. P. S. Kumar, and M. Taj, "Real-time heavy metal detection in water using machine learning-augmented CNT sensors via truncated factorization nuclear norm-based SVD," Microchemical Journal, vol. 218, Art. no. 115375, 2025, doi: 10.1016/j.microc.2025.115375.
[13] M. Yamanouchi, Y. Yokoshiki, M. Dohi, and T. Watanabe, "Machine learning-integrated electrochemical sensors for accurate and continuous free chlorine monitoring," ACS Sensors, vol. 10, no. 11, pp. 8743–8753, 2025, doi: 10.1021/acssensors.5c02634.
[14] F. Abbas, Z. Cai, M. Shoaib, J. Iqbal, M. Ismail, Arifullah, A. F. Alrefaei, and M. F. Albeshr, "Machine learning models for water quality prediction. A comprehensive analysis and uncertainty assessment in Mirpurkhas, Sindh, Pakistan," Water, vol. 16, no. 7, Art. no. 941, 2024, doi: 10.3390/w16070941.
[15] X. Yan, T. Zhang, W. Du, Q. Meng, X. Xu, and X. Zhao, "A comprehensive review of machine learning for water quality prediction over the past five years," Journal of Marine Science and Engineering, vol. 12, no. 1, Art. no. 159, 2024, doi: 10.3390/jmse12010159.
[16] M. Y. Shams, A. A. Elsayed, M. A. Elhoseny, and A. E. Hassanien, "Water quality prediction using machine learning models based on grid search method," Multimedia Tools and Applications, vol. 83, pp. 35307–35334, 2024, doi: 10.1007/s11042-023-16737-4.
[17] K. Karthick, S. Krishnan, and R. Manikandan, "Water quality prediction: A data-driven approach exploiting advanced machine learning algorithms with data augmentation," Journal of Water and Climate Change, vol. 15, no. 2, pp. 431–452, 2024, doi: 10.2166/wcc.2023.403.
[18] M. A. Helaly, S. Rady, M. Mabrouk, M. M. Aref, S. Villarroya, J. M. Cotos, and D. Mera, "Advancements in water quality prediction: A practical review of machine learning and deep learning approaches," Cluster Computing, vol. 28, Art. no. 598, 2025, doi: 10.1007/s10586-025-05221-3.
[19] T. D. Mengistu, I. M. Chung, and S. W. Chang, "Machine learning for water quality prediction and uncertainty assessment," Physics and Chemistry of the Earth, vol. 143, Art. no. 104319, 2026, doi: 10.1016/j.pce.2026.104319.
[20] U. Ahmed, R. Mumtaz, H. Anwar, A. Shah, R. Irfan, and J. García-Nieto, "Efficient water quality prediction using supervised machine learning," Water, vol. 11, no. 11, Art. no. 2210, 2019, doi: 10.3390/w11112210.
[21] K. Chen, "Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data," Water Research, vol. 171, Art. no. 115454, 2020, doi: 10.1016/j.watres.2019.115454.
[22] H. Lu and X. Ma, "Hybrid decision tree-based machine learning models for short-term water quality prediction," Chemosphere, vol. 249, Art. no. 126169, 2020, doi: 10.1016/j.chemosphere.2020.126169.
[23] S. Baek, J. Pyo, and J. Chun, "Prediction of water level and water quality using a CNN LSTM combined deep learning approach," Water, vol. 12, no. 12, Art. no. 3399, 2020, doi: 10.3390/w12123399.
[24] Y. Liu, Y. Liang, K. Ouyang, S. Liu, D. Rosenblum, and Y. Zheng, "Predicting urban water quality with ubiquitous data: A data-driven approach," IEEE Transactions on Big Data, vol. 8, no. 2, pp. 564–578, 2022, doi: 10.1109/TBDATA.2020.2972564.
[25] T. Kavzoglu and A. Sahin, "Water quality prediction in reservoirs using support vector machines," Journal of Hydrology, vol. 568, pp. 405–415, 2019, doi: 10.1016/j.jhydrol.2018.11.038.
[26] A. Altmann, L. Toloşi, O. Sander, and T. Lengauer, "Permutation importance. A corrected feature importance measure," Bioinformatics, vol. 26, no. 10, pp. 1340–1347, 2010, doi: 10.1093/bioinformatics/btq134.
[27] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S. I. Lee, "From local explanations to global understanding with explainable AI for trees," Nature Machine Intelligence, vol. 2, pp. 56–67, 2020, doi: 10.1038/s42256-019-0138-9.
[28] Colabsss, "Electrochemical Carbon Nano-Sensor Dataset. Water Pollution Monitoring Signals and Levels," Kaggle, 2026. [Online]. Available: https://www.kaggle.com/datasets/colabsss/electrochemical-carbon-nano-sensor-dataset. [Accessed: May 11, 2026].
[29] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
[30] T. Chen and C. Guestrin, "XGBoost. A scalable tree boosting system," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[31] C. Cortes and V. Vapnik, “Support-vector networks,” Machine Learning, vol. 20, no. 3, pp. 273–297, 1995, doi: 10.1007/BF00994018.
[32] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” Nature, vol. 323, pp. 533–536, 1986, doi: 10.1038/323533a0.
[33] T. Fawcett, “An introduction to ROC analysis,” Pattern Recognition Letters, vol. 27, no. 8, pp. 861–874, 2006, doi: 10.1016/j.patrec.2005.10.010.
[34] J. Davis and M. Goadrich, “The relationship between precision recall and ROC curves,” in Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, USA, 2006, pp. 233–240, doi: 10.1145/1143844.1143874.
[35] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you. Explaining the predictions of any classifier,” in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 2016, pp. 1135–1144, doi: 10.1145/2939672.2939778.