Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.1, February 2026, pp. 355~364
e-ISSN: 2548-6861

Outperforming DNN Using MLP in Water Quality Assessment for Aquaculture

Mochammad Anshori[1]*, Mufid Musthofa[2]**
* Informatics, Institut Teknologi, Sains, dan Kesehatan RS.DR. Soepraoen Kesdam V/BRW
** Ilmu Komputer, Brawijaya University
moanshori@itsk-soepraoen.ac.id[1], mufidmusthofa@student.ub.ac.id [2]

Article Info

Article history:

Received 2025-11-23
Revised 2025-12-30
Accepted 2026-01-07

Keyword:

Aquaculture,
Backpropagation,
MLP,
Multilayer Perceptron,
Neural Network.

ABSTRACT

Aquaculture production relies heavily on stable water quality conditions, requiring accurate and efficient assessment methods to support early environmental monitoring and sustainable management. Although deep neural network models have been widely applied to water quality classification, their high computational complexity often limits their applicability in real-time and resource-constrained aquaculture systems. This study aims to evaluate whether a systematically optimized Multilayer Perceptron can outperform a reported deep neural network benchmark in aquaculture water quality assessment while maintaining computational efficiency. The study adopts a structured methodology involving dataset characterization, extreme outlier removal, feature normalization, and stratified data partitioning. A single-hidden-layer Multilayer Perceptron is trained using a feedforward backpropagation learning process, with systematic exploration of hidden neuron configurations and training epochs to identify the optimal architecture. Model performance is evaluated using multiple classification metrics, including accuracy, precision, recall, F1-score, confusion matrix analysis, and receiver operating characteristic and precision–recall curves. Results indicate that the optimal Multilayer Perceptron configuration, consisting of 80 hidden neurons and 200 training epochs, achieves an accuracy of 96.62%, surpassing the deep neural network benchmark accuracy of 95.69%. The proposed model demonstrates strong class-level performance, clear separation between water quality categories, stable convergence behavior, and reduced computational overhead compared to deeper architectures. These findings highlight that increasing model depth does not necessarily improve predictive performance for heterogeneous aquaculture datasets. In conclusion, this study provides empirical evidence that a well-optimized shallow neural network can outperform deeper models in aquaculture water quality assessment. The results emphasize the importance of model parsimony and systematic hyperparameter optimization, offering a practical and efficient solution for real-time aquaculture water quality monitoring applications.

I. INTRODUCTION

Water quality is at the forefront of international environmental management initiatives because it is essential to maintaining human health, agricultural production, and industrial development. [1]. Accurate and timely monitoring of water conditions has become increasingly important as water systems continue to be under strain from urbanization, population increase, and industrial activity. Global development agendas like the Sustainable Development Goals (SDGs), where Goal 3 emphasizes the need to ensure human well-being and Goal 14 highlights the protection of aquatic habitats, also reflect this relevance. [2]. In these contexts, the assessment of Water Quality Indices (WQI) has

become a crucial method for figuring out if water is suitable for various ecological and socioeconomic uses. [3]. Water quality is a key factor in determining organism survival, growth rates, and overall production efficiency in aquaculture, an industry that depends on the stability of physicochemical and biological factors. [4]. The deterioration of water quality, particularly in regions impacted by mining, industrial discharge, and intense human activity, is still being documented by numerous environmental studies. This leads to increased pollution levels and ongoing ecological stress. [5], [6]. These difficulties underscore the need for dependable, effective, and data-driven prediction technologies that can spot changes in water conditions before they become harmful occurrences.

To evaluate water quality in aquaculture systems, practitioners rely on a wide range of indicators representing complex environmental dynamics. Core physicochemical parameters include temperature, turbidity, dissolved oxygen (DO), biochemical oxygen demand (BOD), pH, carbon dioxide (CO_2), alkalinity, hardness, calcium, ammonia, nitrite, phosphorus, and hydrogen sulfide (H_2S), along with biological indicators such as plankton abundance [7], [8], [9], [10], [11]. These variables collectively offer a multidimensional representation of environmental states but also introduce analytical challenges due to their nonlinear interactions, differing scales, and susceptibility to seasonal or anthropogenic disturbances. Traditional approaches to water-quality assessment often struggle to model such nonlinear and heterogeneous data effectively. Consequently, machine learning (ML) has been increasingly embraced as a robust analytical paradigm for capturing complex environmental relationships and generating predictive insights that support early intervention and sustainable management practices. Recent developments in ML have expanded the range of available classification and prediction techniques, making it possible to model water-quality dynamics with improved accuracy and interpretability.

Despite these advancements, several limitations remain apparent in the application of machine learning to aquaculture water-quality prediction. A key challenge concerns the reliance on deep neural networks (DNNs), which, although capable of modeling highly nonlinear patterns, typically require significant computational resources, extensive hyperparameter tuning, and long experimentation cycles. For example, prior work employing a DNN with four hidden layers of 17 neurons each reported an accuracy of 95.69% in aquaculture water-quality assessment [12]. While this result demonstrates the feasibility of deep-learning approaches, the computational demands of such architectures limit their practicality for real-time monitoring or deployment in resource-constrained environments, where rapid retraining and efficient energy consumption are essential. Moreover, DNNs often exhibit sensitivity to initialization, optimizer configurations, and layer depth, making them less accessible for practitioners seeking lightweight solutions that maintain high predictive capability without excessive complexity. These issues suggest that simpler alternatives may offer more promising pathways for environmental classification tasks.

In response to these concerns, growing attention has been directed toward the Multilayer Perceptron (MLP), one of the foundational models in artificial neural networks (ANNs). MLPs are characterized by their relatively simple architecture—consisting of an input layer, one or more hidden layers, and an output layer—and their ability to approximate nonlinear functional relationships through layered transformations [13], [14]. Although MLPs are often viewed as less sophisticated than deep models, numerous studies have shown that when appropriately configured, they are capable of achieving high accuracy in classification and prediction tasks across domains such as hydrological forecasting, medical diagnosis [15], and behavioral analysis [16], [17]. Their performance depends heavily on critical hyperparameters, including the number of hidden neurons, the number of layers, and the number of training epochs, each influencing representational capacity, convergence behavior, and computational efficiency [15], [16]. Because these elements determine the balance between overfitting and generalization, tuning them systematically is essential for uncovering optimal architectures tailored to specific datasets.

At the same time, advancements in aquaculture monitoring have introduced a wide variety of machine-learning frameworks beyond MLPs and DNNs. Studies combining ML with IoT-based sensing infrastructures have enabled continuous and intelligent water-quality monitoring systems capable of supporting sustainable aquaculture operations [7], [8]. Other research has focused on field-based assessment of environmental conditions to understand the ecological impacts of aquaculture systems [9], [10]. More sophisticated ML approaches, such as XGBoost optimized through the Honey Badger Algorithm with SHAP and DiCE explanations, have been proposed to enhance interpretability while maintaining high predictive performance [11]. Additional works utilizing Gradient Boosting, Decision Trees, and CatBoost have provided comparative baselines that demonstrate the diversity of available methods and the potential of ensemble-based classifiers [18], [19], [20]. Meanwhile, beyond aquaculture applications, the broader literature reinforces the flexibility and computational advantages of MLPs, highlighting their effective use in various predictive contexts [13], [14], [15], [16], [17]. Collectively, these studies illustrate both the diversity of techniques available for water-quality modeling and the ongoing need to evaluate the trade-offs between model complexity, interpretability, and computational cost.

Although this growing body of work offers valuable insights, significant research gaps persist. First, while deep-learning frameworks have demonstrated competitive accuracy in water-quality classification, there is limited evidence evaluating whether simpler neural architectures can outperform them when optimally tuned. Second, existing MLP-based studies rarely investigate the combined effects of hidden-neuron counts and epoch settings specifically for

aquaculture datasets, which often contain heterogeneous feature distributions and substantial outliers. Third, the literature tends to prioritize accuracy as the primary metric without sufficiently addressing practical considerations such as model parsimony, interpretability, and efficiency—attributes that are critical for real-world system integration. These gaps underscore the need for systematic research that rigorously explores the potential of lightweight neural architectures to achieve or surpass deep-learning benchmarks while reducing computational burdens.

Motivated by these considerations, the present study aims to determine whether a well-designed MLP can outperform the previously reported DNN model in aquaculture water-quality classification. The research introduces a systematic experimental framework that evaluates a range of hidden-neuron configurations (10–100) and epoch settings (100–1000) to identify the architecture that maximizes predictive accuracy, stabilization, and computational efficiency. While previous studies have reported that shallow models may outperform deep architectures on limited tabular datasets, this work contributes novel insights by providing a systematic and domain-specific evaluation within aquaculture water-quality assessment. Unlike prior studies, this research conducts a structured exploration of hidden-neuron and epoch configurations, integrates outlier handling and normalization tailored to heterogeneous environmental indicators, and demonstrates practical efficiency gains relevant to real-time aquaculture monitoring systems. The novelty of this study lies in its empirical demonstration that a simpler, single-hidden-layer MLP can exceed the performance of a deeper four-layer DNN, challenging the common assumption that deeper architectures inherently yield superior results. Furthermore, by presenting a comprehensive analysis that includes accuracy, precision, recall, F1-score, ROC–AUC, and PR-curve evaluation, the study provides a rigorous foundation for assessing the trade-offs between model performance and computational cost. The scope of this research includes data preprocessing, outlier handling, normalization, model training, hyperparameter tuning, and performance evaluation. Through this investigation, the study contributes to the advancement of practical, efficient, and easily deployable machine-learning solutions for aquaculture water-quality assessment.

II. METHODOLOGY

This study employed a structured methodological framework designed to evaluate the performance of a Multilayer Perceptron (MLP) model in classifying aquaculture water quality. The methodology consists of four principal components as shown in Figure 1 below: (1) dataset description, (2) preprocessing procedures, (3) model architecture and training process, and (4) evaluation strategy. Each component is formulated to ensure reproducibility, methodological rigor, and alignment with established machine-learning standards.

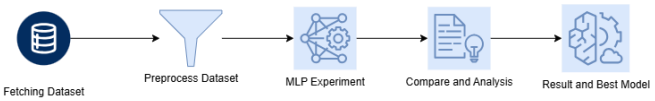

Figure 1. Research Methodology

A. Dataset

The dataset used in this study was obtained from a publicly available Mendeley Data repository (https://data.mendeley.com/datasets/y78ty2g293/1) and consists of 4,300 records representing aquaculture water-quality observations. Each record includes 14 physicochemical and biological features, namely temperature, turbidity, dissolved oxygen, biochemical oxygen demand, carbon dioxide, pH, alkalinity, hardness, calcium, ammonia, nitrite, phosphorus, hydrogen sulfide, and plankton abundance.

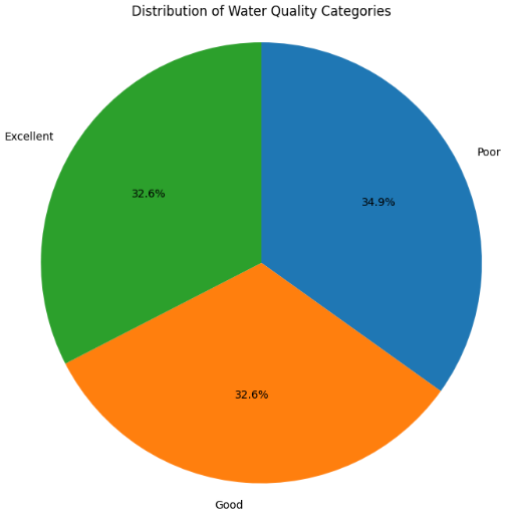

Figure 2. Distribution of water quality class

Figure 2 shows The samples are categorized into three classes: Excellent (1,400 samples, 32.6%), Good (1,400 samples, 32.6%), and Poor (1,500 samples, 34.9%). The dataset is cross-sectional in nature and does not contain explicit temporal or spatial identifiers, which motivates the use of stratified sampling strategies during model evaluation. Fourteen features comprise the input variables: temperature, turbidity, dissolved oxygen (DO), biochemical oxygen demand (BOD), carbon dioxide (CO_2), pH, alkalinity, hardness, calcium, ammonia, nitrite, phosphorus, hydrogen sulfide (H_2S), and plankton abundance. These parameters are widely recognized in the aquaculture domain due to their influence on ecosystem balance and organism health [7]–[11]. However, the dataset presents significant heterogeneity across its variables, including widely varying numeric scales and the presence of extreme values, conditions that necessitate careful preprocessing to ensure stable and unbiased model training [7]–[11].

B. Data Preprocessing

Given the importance of data quality in machine-learning applications, preprocessing was conducted through a structured sequence consisting of outlier handling, normalization, and data partitioning.

1) *Outlier Detection and Removal*: Outlier identification and removal constituted the first stage of preprocessing. Outliers in environmental datasets often arise from instrument errors, sampling inconsistencies, or naturally occurring variations [21]. Their presence may distort statistical relationships, mislead gradient-based learning algorithms, and reduce model generalization capability. Accordingly, this study applied systematic outlier removal procedures recommended in contemporary data-mining literature to minimize noise while preserving essential data characteristics.

2) *Feature Normalization:* Following outlier removal, the remaining features were standardized using z-score normalization, which transforms each feature according to Eq. (1).

$$z = \frac{x-\mu}{\sigma} \quad (1)$$

where x is the original value, μ is the mean, and $\sigma$is the standard deviation [22]. This method was chosen due to the substantial disparities in feature scales. For example, plankton counts range into the thousands, while contaminants such as ammonia and H_2S appear in near-zero ranges. Standardization ensures proportional influence of all features during weight updates and improves convergence stability during training where x denotes the original feature value, $\mu$the mean, and $\sigma$the standard deviation [22].

3) *Stratified Training-Testing Split:* The dataset was subsequently divided into training and testing subsets using stratified holdout sampling, preserving class proportions across partitions [23]. To align with common practice in machine learning, an 80:20 split was applied, following established recommendations that balance generalization evaluation with training sufficiency [24], [25]. Considering the distributional changes brought forth by outlier removal, stratification was very crucial. This situation is based on the Pareto principle, which is the best approach in a lot of studies [26], [27].

C. MLP Model Architecture and Training Procedure

The core methodological component involved desgining and training the MLP architecture while systematically testing the influence of different hyperparameter configurations.

1) *Model Architecture*: The MLP foundation model is shown at Figure 3 below. The MLP model comprises three layers: an input layer, a single hidden layer, and an output layer [14]. The MLP model used in this research consists of three layers: an input layer with 14 normalized features, a single hidden layer, and an output layer for producing predictions across the three water-quality classes. The structure reflects the architecture commonly associated with classical feedforward neural networks.

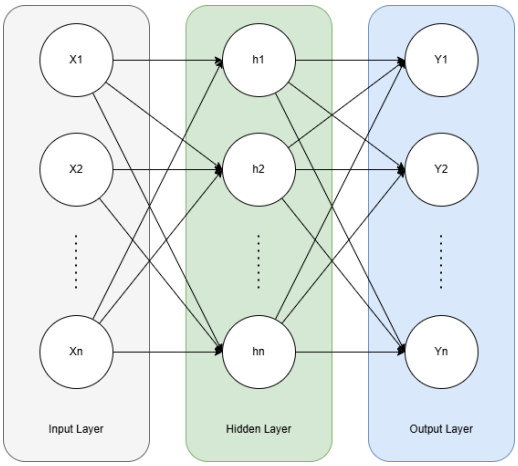

Figure 3. The illustration of MLP architecture foundation.

To evaluate the effects of architectural capacity, the number of neurons in the hidden layer was varied. Each hidden neuron performs a nonlinear transformation of the weighted input signals, guided by the ReLU activation function—selected for its computational efficiency and ability to mitigate vanishing gradients [26], [27]. The ReLU function follows Eq. (2):

$$feedforward = max\left(0, \sum_i w_{ij} X_i + b_i\right) \quad (2)$$

2) *Training Algorithm:* Training was Training was conducted via the standard feedforward–backpropagation procedure. In feedforward propagation, weighted inputs are transformed through the activation function to produce class logits. Backpropagation then computes gradients of the loss function with respect to every weight and bias. To ensure effective parameter updates, the Adam optimizer was employed, consistent with its documented advantages in adaptability and convergence efficiency [28]. Parameter updates follow Eq. (3):

$$\theta_t = \theta_{t-1} - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \quad (3)$$

where $\hat{m}_t$ and $\hat{v}_t$ denote bias-corrected first and second moment estimates, respectively.

3) *Hyperparameter Tuning Framework*: Because MLP performance depends critically on hyperparameters [15], [16], the study implemented a structured experiment to determine the best MLP model. Two hyperparameter dimensions were explored:

- Hidden neurons: 10 to 100 (step size: 10)
- Epochs: 100 to 1000 (step size: 100)

This approach ensured balanced exploration of model complexity and learning duration, enabling identification of

an architecture that maximizes accuracy while minimizing computational overhead. The systematic combination of these parameters allowed the study to evaluate trends in convergence, generalization, and training efficiency.

D. Evaluation Metrics and Analytical Procedures

Model performance was assessed using a comprehensive set of widely accepted classification metrics to ensure robust evaluation across classes.

1) *Accuracy*: Overall classification accuracy was computed using Eq. (4)

$$accuracy = \frac{TP + TN}{N} \quad (4)$$

where TP denotes true positives, TNtrue negatives, and $N$the total number of samples.

2) *Precision, Recall and F1-score:* Precision, recall, and F1-score were calculated to assess class-level performance, especially important due to class imbalance after outlier removal. The metrics follow Eqs. (5)-(7):

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

$$F1 - score = 2 \cdot \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

These indicators help evaluate the model's sensitivity to false positives and false negatives, a critical requirement for environmental classification tasks [29], [30].

3) *ROC-AUC and Precision-Recall Curves*: To evaluate discriminatory ability across threshold settings. Receiver Operating Characteristic (ROC) curves and the Area Under the Curve (AUC) were generated for all three classes. AUC values approaching 1.0 signify strong separability [31], [32]. Precision–Recall (PR) curves were also analyzed due to their suitability for imbalanced datasets [33].

4) *Confusion Matrix Visualization*: Confusion matrices were computed to quantify misclassification and identify class paris with overlapping feature space. This visualization supports diagnostic interpretation of model behavior and error patterns [34].

III. Results And Discussion

The results of this study provide a comprehensive evaluation of the performance of the Multilayer Perceptron (MLP) model in classifying aquaculture water quality and demonstrate how an optimized MLP configuration can surpass the accuracy achieved by a previously reported Deep Neural Network (DNN). The findings are presented through a systematic analysis encompassing descriptive statistics, effects of preprocessing, hyperparameter exploration, computational efficiency, and performance interpretation using multiple evaluation metrics. Together, these analyses offer compelling evidence supporting the viability of a parsimonious neural architecture for environmental classification tasks.

A. Descriptive Statistical Analysis of Water Quality Features

Figure 4. Dataset Descriptive

Initial descriptive analysis revealed substantial heterogeneity across the fourteen physicochemical and biological indicators included in the dataset. The descriptive heatmap showed that features such as plankton abundance exhibit extremely wide ranges—for instance, values span from approximately 78.60 to 7,460.42—while others, such as ammonia and H_2S, cluster near zero with narrow distributions. This discrepancy in scale indicates the intrinsic complexity of environmental datasets and underscores the necessity of normalization to prevent high-magnitude features from dominating gradient-based optimization during model training. Without such preprocessing, neural networks may experience unstable convergence or biased learning trajectories.

B. Outlier Removal

The violin plots further illustrated the presence of extreme values in several indicators as shown at Figure 5. Particularly ammonia, nitrite, and H_2S. These elongated density shapes suggest that a significant number of observations fall far from the main distribution body. As documented in the literature, such outliers can compromise the generalization capability of machine-learning models by introducing noise and distorting feature–label relationships during training. Accordingly, systematic outlier removal was implemented as a critical preprocessing step.

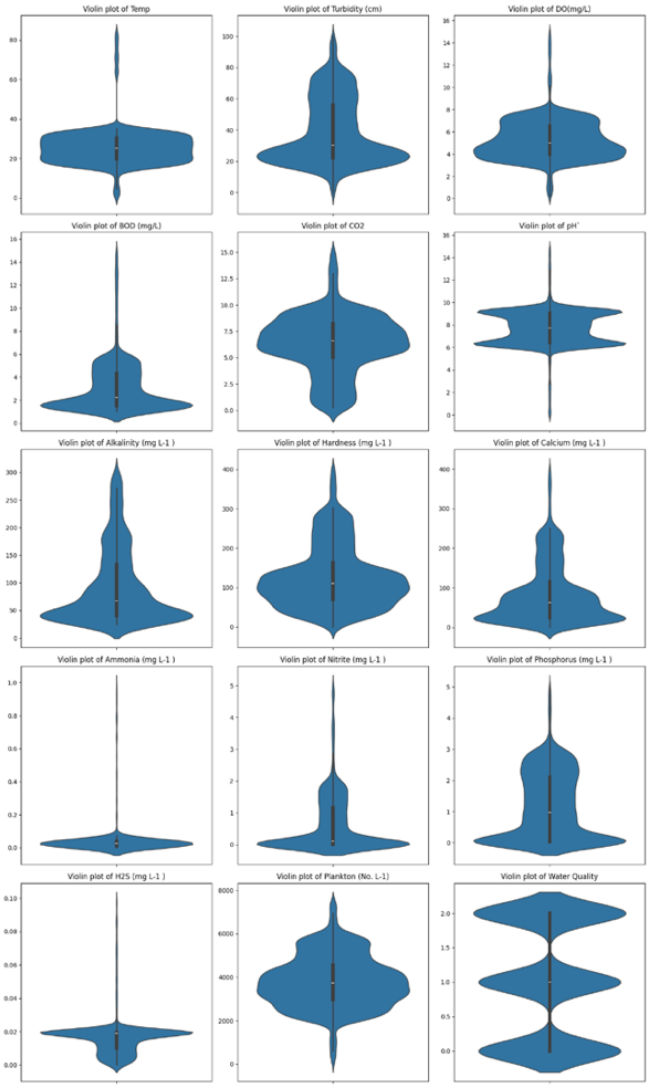

Figure 5. Violin plot of dataset to check for outliers

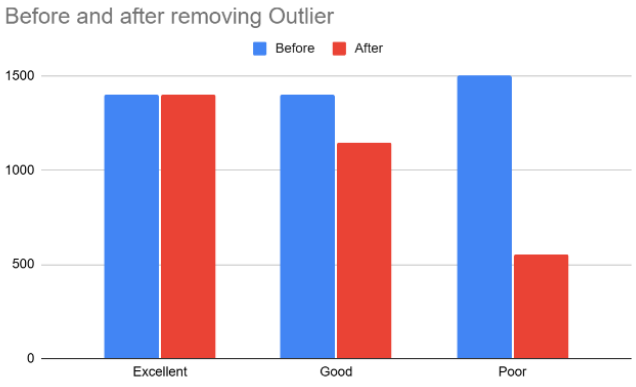

Figure 6. Dataset distribution after removing outlier

Shown at Figure 6, outlier removal substantially altered the class distribution within the dataset. While the Excellent class remained largely stable at around 1,400 samples, the Poor class experienced a dramatic reduction from 1,500 samples to approximately 500. This result suggests that many Poor-class observations exhibited extreme or atypical feature values. The resulting imbalance has implications for classification difficulty, as boundary distinction between the Good and Poor classes becomes more challenging when sample representation is unequal. This effect is reflected in certain misclassification patterns observed later in the analysis. Nonetheless, removing outliers improved feature homogeneity and contributed to more reliable training behavior.

C. Hyperparameter Analysis: Accuracy Across MLP Configurations

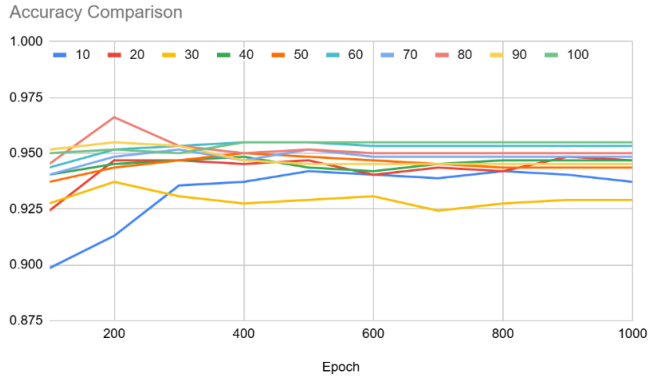

Figure 7. Accuracy comparison on each parameter test

A core component of the investigation involved assessing how variations in hidden-neuron counts and training epochs influence MLP performance. According to Figure 7, accuracy trends demonstrated that the model reached strong classification performance even at early stages of learning, surpassing 0.94 after approximately 400 epochs in most configurations. This rapid convergence reflects the separability of the dataset and the MLP's inherent capability to model nonlinear relationships in water-quality indicators. At 200 epochs, several noteworthy patterns emerged. The configuration with 80 hidden neurons yielded an accuracy of approximately 0.95, while the configuration with 20 hidden neurons momentarily achieved an accuracy near 0.9662 before exhibiting stagnation suggestive of convergence to a suboptimal local minimum. By contrast, the 80-neuron model maintained stable performance across training durations, indicating a more robust representational capacity. In general, smaller models occasionally produced high early accuracy but failed to maintain consistent performance with further training, whereas larger models demonstrated smoother accuracy trajectories but incurred higher computational costs. These findings illustrate the importance of balancing model complexity with adaptive capacity.

D. Optimal Model and Computational Efficiency

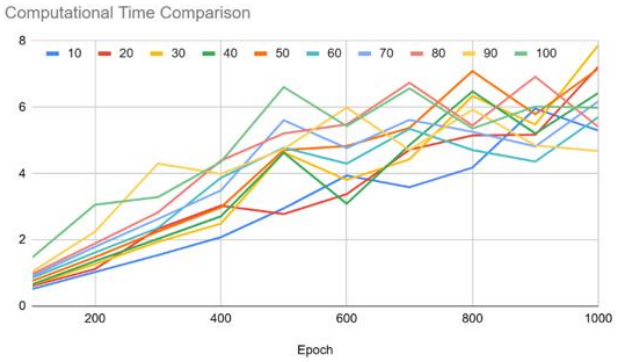

Figure 8. Computational time in each experiment

Comparative analysis across all hyperparameter combinations identified the optimal architecture: a single hidden layer with 80 neurons trained over 200 epochs, achieving an accuracy of 96.62%. This result exceeds the previously reported DNN accuracy of 95.69%, providing strong evidence that deeper architectures are not inherently superior for this task. Computational efficiency was also examined. The optimal model required only 1.89 units of computation time, significantly less than larger models trained for extended epochs. Models with 100 neurons trained for 600–1000 epochs exhibited no improvement in accuracy beyond approximately 0.9549, yet incurred training times exceeding 5–6 units. This demonstrates diminishing returns in neural-network complexity: increasing the number of neurons or epochs beyond certain thresholds does not enhance predictive capability but instead introduces unnecessary computational overhead. The optimal configuration thus aligns with principles of Pareto optimality: minimal complexity for maximal performance.

E. Final MLP Architecture

TABEL I
COMPARISON MODEL IN EACH EPOCH

Epoch	Hidden Neuron	Accuracy	Execution Time (second)
100	90	0.9516908213	1.04
200	80	0.9661835749	1.89
300	60	0.9533011272	2.36
400	60	0.9549114332	3.87
500	60	0.9549114332	4.78
600	100	0.9549114332	5.43
700	100	0.9549114332	6.57
800	100	0.9549114332	5.36
900	100	0.9549114332	6.02
1000	100	0.9549114332	5.99

The Table 1 decisively confirms the superiority of the proposed method (MLP with a single hidden layer, 80 neurons, and 200 epochs) over the prior approach (Deep Neural Network with 4 hidden layers and 17 neurons). The proposed MLP achieved an accuracy of 96.62%, representing a significant increase of 0.93% compared to the prior DNN's 95.69%. Academically, this enhanced performance, despite using a simpler, shallower architecture, argues that precise hyperparameter optimization (80 neurons, 200 epochs) was more effective in accurately classifying water quality features than merely increasing the network's depth.

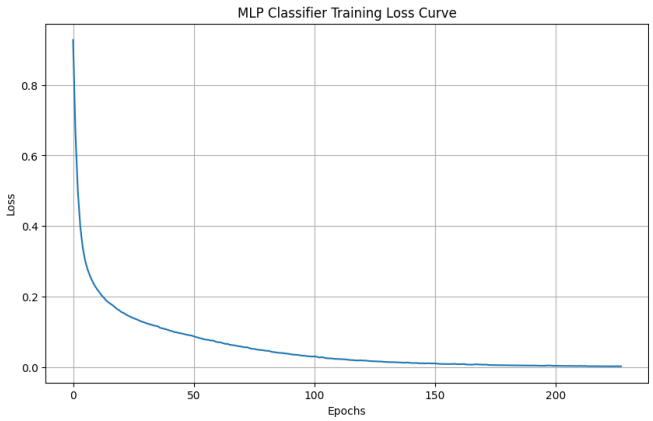

Figure 9. The Classifier training loss curve

The presented Figure 9 illustrates the training loss trajectory of a Multilayer Perceptron (MLP) classifier, which is characterized by a sharp initial decline from a loss value of approximately 0.9. Subsequently, the rate of descent diminishes significantly, with the curve demonstrating a gradual convergence towards a near-zero loss value as the number of epochs exceeds 150.

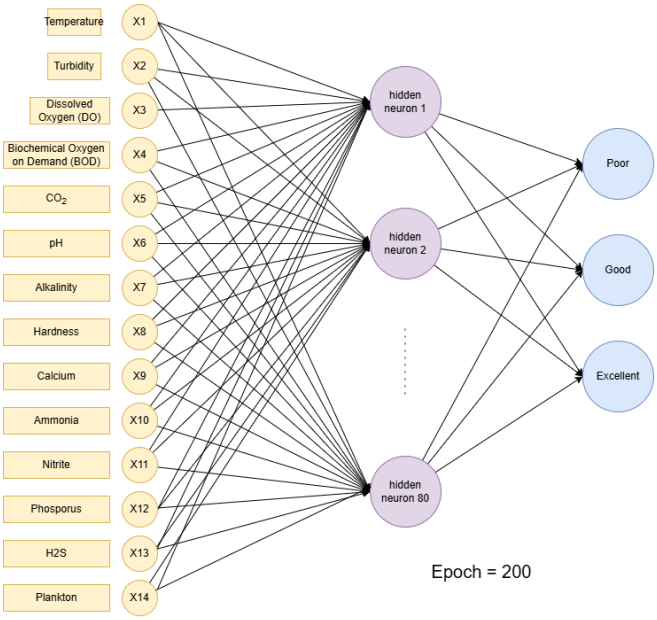

Figure 10. The best MLP model with hidden neuron = 80 and epoch = 200

The optimal model comprises 14 input features feeding into a single hidden layer of 80 neurons, followed by a three-

node output layer corresponding to the Poor, Good, and Excellent water-quality classes shown at Figure 10. This architecture achieves a balance between expressive capacity and computational simplicity, reinforcing the broader premise that shallow networks can outperform deeper ones when hyperparameters are systematically optimized.

F. Confusion Matrix Analysis

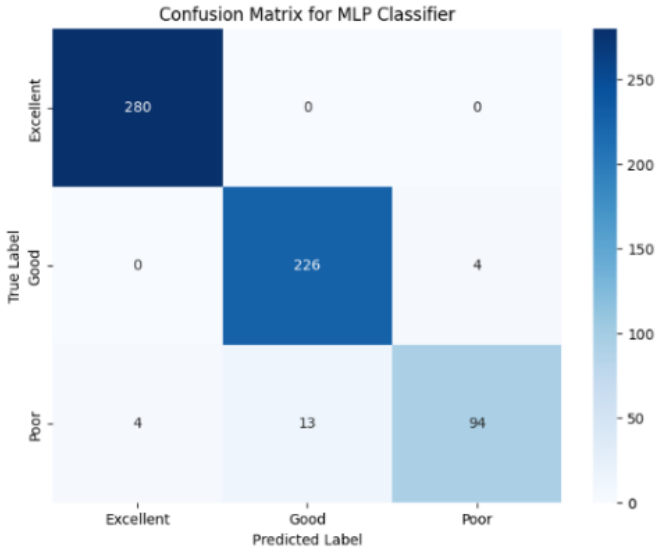

Figure 11. Confusion matrix of the model

The confusion matrix for the optimal MLP model indicates excellent classification performance across all classes represent in Figure 11. The Excellent class achieved 100% correct classification, reflecting strong separability in feature space. The Good and Poor classes also attained high accuracy, with the majority of samples assigned correctly (226 Good, 94 Poor). Misclassifications primarily occurred between the Good and Poor classes, consistent with earlier observations about distribution shifts following outlier removal. Specifically, 4 Good samples were misclassified as Poor, while 13 Poor samples were misclassified as Good. A small number of Poor samples (4) were misclassified as Excellent. Despite these discrepancies, total misclassifications remained low (17), demonstrating robust overall performance

G. ROC-AUC And Precision-Recall Performance

The Receiver Operating Characteristic (ROC) curves and their corresponding Area Under the Curve (AUC) scores were used to assess the model's discriminative power, with the results confirming its effectiveness in distinguishing between classes across various threshold settings. AUC values were exceptionally high, registering 1.00 for the Excellent class, 0.99 for Good, and 0.97 for Poor.

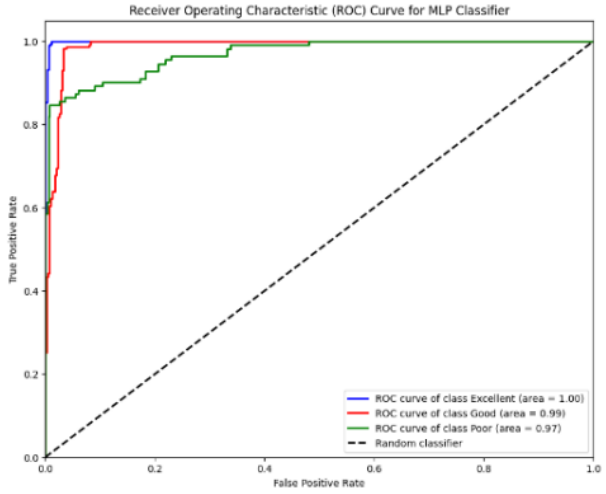

Figure 12. ROC and AUC score

While the Poor class exhibited a slightly lower AUC score, likely due to class imbalance and reduced sample representation, the value of 0.97 remains strong enough to indicate reliable separability between the classes.

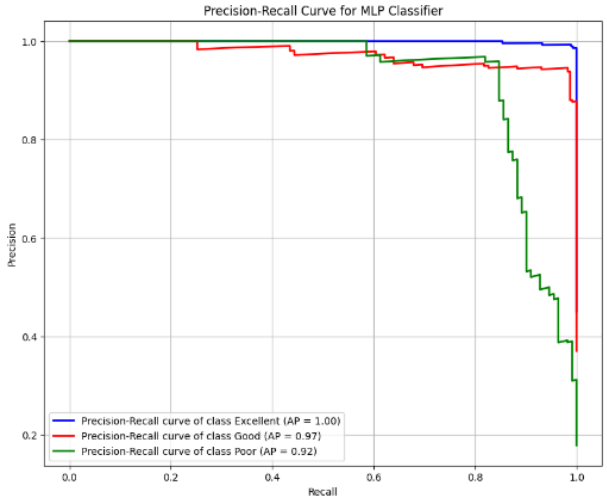

Figure 13. PR-Curve of the Model

A Precision-Recall (PR) analysis was performed to evaluate the model's performance specifically under conditions of class imbalance, yielding high Average Precision (AP) values. The AP scores were 1.00 for the Excellent class, 0.97 for Good, and 0.92 for Poor. The PR curves for the Excellent and Good classes remained near the upper region across all recall ranges, demonstrating consistent high precision. Although the Poor class curve showed an expected decline in precision at extreme recall values due to its reduced sample representation, its AP score of 0.92 still signifies strong performance for practical classification scenarios.

H. Comparative Evaluation and Discussion

TABEL 2
COMPARISON MODEL IN EACH EPOCH

Method	Parameters	Accuracy
Deep Neural Network: [12]	Hidden layer = 4; hidden neurons = 17	95.69%
Proposed method using MLP	Hidden neurons = 80; epoch = 200	96.62%

Table 2 shows the comparison between prior research and this research. The optimized Multilayer Perceptron (MLP) configuration demonstrated clear superiority over the previously reported Deep Neural Network (DNN) benchmark. The Deep Neural Network (DNN) used as a benchmark in this study follows the architecture reported in [12] which consists of four hidden layers with 17 neurons per layer and was trained using a supervised learning framework on the same dataset. The objective of this study is not to redesign or further optimize the DNN architecture, but to evaluate whether a systematically tuned shallow MLP can surpass a commonly referenced deep-learning baseline under comparable data conditions. Specifically, the proposed single-hidden-layer MLP achieved 96.62% accuracy, surpassing the 95.69% accuracy of the DNN, which utilized four hidden layers of 17 neurons each. This improvement is significant both numerically and conceptually, reinforcing that model depth does not necessarily correlate with predictive performance in environmental classification tasks.

The overall findings underscore several key insights: effective preprocessing (normalization and outlier removal) is essential for stabilizing training on heterogeneous environmental datasets; model parsimony (minimal complexity with strong performance) is both attainable and advantageous in real-world aquaculture monitoring; hyperparameter optimization is more critical than model depth, particularly in MLP architectures; and generalization performance is robust across all water-quality classes, as confirmed by ROC-AUC and Precision-Recall curves. Collectively, the results establish that a carefully tuned MLP can surpass deeper models while maintaining computational efficiency, making it a highly compelling solution for real-time and resource-constrained aquaculture water-quality assessment applications.

IV. CONCLUSION

This study demonstrates that a carefully configured Multilayer Perceptron can surpass the performance of a previously reported Deep Neural Network in aquaculture water-quality classification while substantially reducing computational demands. Through systematic experimentation on hidden-neuron sizes and training epochs, the optimal MLP architecture comprising a single hidden layer with 80 neurons trained for 200 epochs achieved an accuracy of 96.62%, exceeding the DNN benchmark of 95.69%. The model also exhibited strong class separability, high precision across all categories, and stable learning behavior, despite challenges introduced by heterogeneous feature scales and class imbalance after outlier removal. These findings indicate that architectural depth is not always a prerequisite for achieving high predictive capability. Instead, model parsimony, supported by rigorous hyperparameter optimization, can provide an effective and computationally efficient alternative for environmental classification tasks. This contributes to current knowledge by offering empirical evidence that shallow neural architectures can outperform deeper models when aligned with the characteristics of the dataset. Future work may explore automated hyperparameter-search methods, integration with real-time sensor data, and validation across more diverse aquaculture environments. Such extensions would further establish the practical utility of lightweight neural models for sustainable and responsive water-quality monitoring.

REFERENCES

[1] N. Luvhimbi, T. G. Tshitangano, J. T. Mabunda, F. C. Olaniyi, and J. N. Edokpayi, "Water quality assessment and evaluation of human health risk of drinking water from source to point of use at Thulamela municipality , Limpopo Province," *Sci Rep*, no. 0123456789, pp. 1–17, 2022, doi: 10.1038/s41598-022-10092-4.

[2] P. Ghosh and T. Chakravarty, "Sustainable Development Goals: Challenges, Opportunities, and the Way Forward," in *Sustainability: Science, Policy, and Practice in India: Challenges and Opportunities*, V. Dutta and P. Ghosh, Eds., Cham: Springer International Publishing, 2023, pp. 235–241. doi: 10.1007/978-3-031-50132-6_17.

[3] B. L. Widiyanti, "Assessment of water quality and its effect on the health of Selong District Resident, East Lombok, West Nusa Tenggara Province," *IOP Conf Ser Earth Environ Sci*, vol. 986, no. 1, p. 12078, Feb. 2022, doi: 10.1088/1755-1315/986/1/012078.

[4] I. Effendi, "Pengembangan Akuakultur pada Lahan Suboptimal Menuju," *Prosiding Seminar Nasional Lahan Suboptimal*, vol. 7, no. 1, pp. 978–979, 2019.

[5] R. Nawaz *et al.*, "Water Quality Index and Human Health Risk Assessment of Drinking Water in Selected Urban Areas of a Mega City," 2023.

[6] A. Mohammadpour, E. Gharehchahi, and M. Amiri, "Assessment of drinking water quality and identifying pollution sources in a chromite mining region," *J Hazard Mater*, vol. 480, no. June, p. 136050, 2024, doi: 10.1016/j.jhazmat.2024.136050.

[7] A. Al, M. Hridoy, C. Bordin, A. Masood, and K. Masood, "Results in Chemistry Predictive modelling of aquaculture water quality using IoT and advanced machine learning algorithms," *Results Chem*, vol. 16, no. June, p. 102456, 2025, doi: 10.1016/j.rechem.2025.102456.

[8] R. Baena-navarro, Y. Carriazo-regino, F. Torres-hoyos, and J. Pinedo-l, "Intelligent Prediction and Continuous Monitoring of Water Quality in Aquaculture : Integration of Machine Learning and Internet of Things for Sustainable Management," 2025.

[9] N. A. Regency, R. Ezraneti, and M. Khalil, "Study of Water Quality for Aquaculture System in Ujung Pacu River, North Aceh Regency," *Omni-Akuatika*, vol. 15, no. 3, pp. 119–127, 2020, doi: http://dx.doi.org/10.20884/1.oa.2020.16.3.849.

[10] R. P. Wisnu, M. Karuniasa, and S. S. Moersidik, "The effect of fish aquaculture on water quality in Lake Cilala, Bogor Regency," *IOP Conf Ser Earth Environ Sci*, vol. 399, no. 1, p. 12111, Dec. 2019, doi: 10.1088/1755-1315/399/1/012111.

[11] S. M. Naim, P. Das, and J. Tiang, "Aquaculture Water Quality Classification Using XGBoost Classifier Model Optimized by the

Honey Badger Algorithm with SHAP and DiCE-Based Explanations," pp. 1–18, 2025.

[12] R. Arabelli, T. Bernatin, and V. Veeramsetty, "Water quality assessment for aquaculture using deep neural network," *Desalination Water Treat*, vol. 321, no. October 2024, p. 101016, 2025, doi: 10.1016/j.dwt.2025.101016.

[13] A. Lazcano and M. A. Jaramillo-mor, "Back to Basics : The Power of the Multilayer Perceptron in Financial Time Series Forecasting," *Mathematics*, vol. 12, no. 12, pp. 1–18, 2024, doi: 10.3390/math12121920.

[14] Enung, H. Kasyanto, and R. R. Sari, "Penerapan Algoritma Multilayer Perceptron (MLP) Untuk Memprediksi Debit Di Sungai Citarum Bagian Hulu (Pos Pengukuran Majalaya), Kab.Bandung, Jawa Barat," *Potensi: Jurnal SIpil Politeknik*, vol. 25, no. 1, pp. 1–8, 2023, doi: https://doi.org/10.35313/potensi.v25i1.4513.

[15] A. D. Novika and A. S. Girsang, "Multi-layer perceptron hyperparameter optimization using Jaya algorithm for disease classification," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 35, no. 1, pp. 620–630, 2024, doi: 10.11591/ijeecs.v35.i1.pp620-630.

[16] K. Amadea, F. A. Bachtiar, and G. Pangestu, "Klasifikasi Pola Pergerakan Bola Mata Menggunakan Metode Multilayer Backpropagation," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 9, no. 2, pp. 391–400, 2022, doi: 10.25126/jtiik.202295668.

[17] M. Anshori, M. S. Haris, and W. Teja Kusuma, "Penerapan Backpropagation Neural Network (BPNN) Untuk Prediksi Kecanduan Smartphone Pada Remaja," *Cices*, vol. 9, no. 2, pp. 192–202, 2023, doi: 10.33050/cices.v9i2.2701.

[18] N. Nasir *et al.*, "Water quality classification using machine learning algorithms," *Journal of Water Process Engineering*, vol. 48, p. 102920, 2022, doi: https://doi.org/10.1016/j.jwpe.2022.102920.

[19] U. Machine and L. Techniques, "Prediction of Water Quality Classification of the Kelantan River Basin , Malaysia , Using Machine Learning Techniques," *Water (Basel)*, vol. 14, no. 7, 2022, doi: 10.3390/w14071067.

[20] J. Yung *et al.*, "Towards a time and cost effective approach to water quality index class prediction," *J Hydrol (Amst)*, vol. 575, no. March, pp. 148–165, 2019, doi: 10.1016/j.jhydrol.2019.05.016.

[21] C. Sanjeev, K. Dash, A. Kumar, S. Dehuri, and A. Ghosh, "An outliers detection and elimination framework in classification task of data mining," *Decision Analytics Journal*, vol. 6, no. January, p. 100164, 2023, doi: 10.1016/j.dajour.2023.100164.

[22] O. Colliot, *A Non-technical Introduction to Machine Learning*, vol. 197. Paris: Humana Press, 2023. doi: 10.1007/978-1-0716-3195-9_1.

[23] F. Farias, T. B. Ludermir, and C. Bastos-filho, "Similarity Based Stratified Splitting : an approach to train better classifiers," no. November, 2020, doi: 10.48550/arXiv.2010.06099.

[24] A. Bhaskara, A. A. Ginting, and A. M. Masagala, "Penerapan Konstruksi Ramping terhadap Waste pada Ruang Lingkup Manajemen Proyek (Studi Kasus: Proyek Pembangunan Gedung Auditorium di Yogyakarta)," *Semesta Teknika*, vol. 25, no. 1, pp. 80–88, 2022, doi: 10.18196/st.v25i1.13403.

[25] F. N. Arifah *et al.*, *Machine Learning: Seni dan Sains Belajar dari Data*. Yayasan Kita Menulis, 2025.

[26] M. Li, "Comprehensive Review of Backpropagation Neural Networks," *Academic Journal of Science and Technology*, vol. 9, no. 1, 2024, doi: https://doi.org/10.54097/51y16r47.

[27] S. H. Gulo, A. H. Lubis, T. Informatika, F. Teknik, and U. M. Area, "Penerapan Multi-Layer Perceptron untuk Mengklasifikasi Penduduk Kurang Mampu," *EXPLORER Journal of Computer Science and Information Technology*, vol. 4, no. 2, pp. 51–59, 2024, doi: 10.47065/explorer.v4i2.1146.

[28] L. Yang, "Theoretical Analysis of Adam Optimizer in the Presence of Gradient Skewness," *International Journal of Applied Science*, vol. 7, no. 2, pp. 27–42, 2024, doi: https://doi.org/10.30560/ijas.v7n2p27.

[29] M. Musthofa and M. Anshori, "Comparing Discriminant Analysis Function for Early Prediction of Smartphone Addiction," *Journal of Enhanced Studies in Informatics and Computer Applications*, vol. 2, no. 1, pp. 1–7, 2025, doi: https://doi.org/10.47794/jesica.v2i1.12.

[30] M. Anshori, W. T. Kusuma, and R. S. Pradini, "Prediction of Sleep Disorder : Insomnia Using AdaBoost Ensemble Learning Algorithm with Grid Search Optimization," *InComTech: Jurnal Telekomunikasi dan Komputer*, vol. 14, no. 1, pp. 30–38, 2024, doi: 10.22441/incomtech.v14i1.19306.

[31] Muttaqin *et al.*, *Pengenalan Data Mining*, First Edit., no. July. Yayasan Kita Menulis, 2023.

[32] M. C. Saletia, M. Anshori, and S. Haris, "Comparing Different KNN Parameters Based on Woman Risk Factors to Predict the Cervical Cancer," *Journal of Applied Informatics and Computing (JAIC)*, vol. 9, no. 5, p. 2179, Oct. 2025, doi: https://doi.org/10.30871/jaic.v9i5.10746.

[33] J. Cook, "When to consult precision-recall curves," *Stata J*, vol. 20, no. 1, pp. 131–148, 2020, doi: 10.1177/1536867X20909693.

[34] M. Anshori and M. S. Haris, "Predicting Heart Disease using Logistic Regression," *Knowledge Engineering and Data Science*, vol. 5, no. 2, p. 188, 2022, doi: 10.17977/um018v5i22022p188-196.