International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

# A MACHINE LEARNING APPROACH FOR EARLY DETECTION OF FISH DISEASES BY ANALYSING WATER QUALITY

J. KUMARI[1], CH.VIJAYA ANANDA LAKSHMI[2]
ASSISTANT PROFESSOR[1], PG SCHOLAR[2]
DEPARTMENT OF MASTER OF COMPUTER APPLICATIONS
QIS COLLEGE OF ENGINEERING & TECHNOLOGY,
VENGAMUKKAPALEM (V), ONGOLE,
PRAKASAM DIST, ANDHRA PRADESH.

## ABSTRACT

Aquaculture plays a vital role in global food production, but fish diseases pose a significant threat to productivity and sustainability. Early detection of fish diseases is crucial to prevent large-scale losses and ensure healthy aquatic ecosystems. This project proposes a machine learning-based system that analyzes water quality parameters to predict and detect fish diseases at an early stage. The system collects real-time data such as pH, temperature, dissolved oxygen, ammonia, and turbidity using sensors. Machine learning models including Random Forest, Support Vector Machine, and Neural Networks are employed to analyze patterns and identify anomalies associated with disease outbreaks. By correlating water quality variations with fish health conditions, the system provides timely alerts and actionable insights to fish farmers. The proposed approach enhances disease prediction accuracy, reduces manual monitoring efforts, and promotes sustainable aquaculture practices.

## INTRODUCTION

Aquaculture has emerged as one of the fastest-growing sectors in food production, contributing significantly to global nutrition and economic development. However, fish diseases remain a major challenge, causing substantial losses to farmers and affecting overall productivity. These diseases are often influenced by water quality parameters such as pH, temperature, dissolved oxygen, ammonia concentration, and turbidity. Traditional methods of disease detection rely on manual observation of fish behavior and physical symptoms, which are often identified at later stages when the disease has already spread. This delay in detection leads to increased mortality rates and economic losses.

With advancements in technology, machine learning has become a powerful tool for analyzing complex environmental data and identifying patterns that are not easily observable by humans. By integrating sensor-based data

Received: 10-05-2026 | Accepted: 18-06-2026 | Published: 24-06-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

collection with machine learning algorithms, it is possible to continuously monitor water quality and detect early signs of disease outbreaks. Models such as Random Forest, Support Vector Machine, and Neural Networks can process large datasets and provide accurate predictions based on historical and real-time data. This project aims to develop an intelligent system that monitors water quality parameters and predicts fish diseases at an early stage, enabling timely intervention and improved aquaculture management. The system not only enhances productivity but also contributes to sustainable and efficient fish farming practices

## SYSTEM ARCHITECTURE

The proposed system uses IoT sensors and machine learning techniques to monitor water

quality and predict fish diseases at an early stage. Sensors collect parameters such as pH, temperature, dissolved oxygen, ammonia, and turbidity. The collected data is preprocessed and analyzed using machine learning models like Random Forest, Support Vector Machine (SVM), and Neural Networks. Based on the prediction results, the system generates alerts to help fish farmers take preventive actions and


maint.

## LITERATURE SURVEY

### 1. Title: Machine Learning for Fish Disease Detection

**Authors:** S. Kumar et al. (2020)
**Merits:** High prediction accuracy
**Demerits:** Requires large datasets

### 2. Title: Water Quality Monitoring in Aquaculture

**Authors:** R. Singh et al. (2019)
**Merits:** Real-time monitoring
**Demerits:** Sensor maintenance issues

### 3. Title: AI-Based Aquaculture Management System

**Authors:** P. Sharma et al. (2021)
**Merits:** Automated decision-making
**Demerits:** High implementation cost

### 4. Title: Fish Disease Prediction Using Neural Networks

**Authors:** A. Verma et al. (2018)
**Merits:** Captures complex patterns
**Demerits:** Computationally expensive

Received: 10-05-2026 | Accepted: 18-06-2026 | Published: 24-06-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

**5. Title: IoT-Based Smart Fish Farming**

**Authors:** K. Patel et al. (2022)
**Merits:** Remote monitoring
**Demerits:** Network dependency

**6. Title: Support Vector Machine for Water Quality Analysis**

**Authors:** L. Wang et al. (2017)
**Merits:** High classification accuracy
**Demerits:** Sensitive to parameter tuning

**7. Title: Random Forest for Environmental Monitoring**

**Authors:** J. Lee et al. (2019) **Merits:** Robust and efficient
**Demerits:** Less interpretable

**8. Title: Deep Learning for Aquaculture Disease Detection**

**Authors:** M. Reddy et al. (2023)
**Merits:** High accuracy **Demerits:** Requires high computational power

**9. Title: Sensor-Based Water Quality Analysis**

**Authors:** S. Verma et al. (2020)
**Merits:** Continuous monitoring
**Demerits:** Hardware cost

**10. Title: Predictive Analytics in Aquaculture**

**Authors:** A. Jain et al. (2021)
**Merits:** Early detection capability **Demerits:** Data preprocessing complexity

## SYSTEM ANALYSIS

### EXISTING SYSTEM

The existing systems for fish disease detection primarily rely on manual observation and traditional water quality testing methods. Farmers typically monitor fish behavior, feeding patterns, and physical symptoms to identify diseases, which often leads to late detection. Laboratory testing of water samples is also used to assess water quality; however, this process is timeconsuming, expensive, and does not provide real-time insights. Some modern systems incorporate basic sensors to monitor parameters such as temperature and pH, but they lack advanced analytical capabilities and predictive features. Additionally, these systems do not effectively correlate water quality data with fish health conditions, limiting their usefulness in early disease detection. The absence of automated analysis and intelligent decision-making results in delayed responses and increased risk of disease spread. Therefore, existing systems are insufficient for proactive aquaculture management and require significant improvements.

### DISADVANTAGES

- Manual and delayed detection
- Lack of real-time monitoring
- No predictive capability

Received: 10-05-2026 | Accepted: 18-06-2026 | Published: 24-06-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

- High operational cost
- Limited scalability
- Inefficient disease management

## PROPOSED SYSTEM

The proposed system introduces a machine learning-based approach for early detection of fish diseases by analyzing water quality parameters in real time. The system integrates IoT sensors to continuously monitor key parameters such as pH, temperature, dissolved oxygen, ammonia, and turbidity. The collected data is transmitted to a

centralized system where it is processed and analyzed using machine learning algorithms such as Random Forest, Support Vector Machine, and Neural Networks. These models are trained on historical data to identify patterns and correlations between water quality variations and fish disease occurrences.

The system is capable of detecting anomalies in water conditions that may indicate the onset of diseases, enabling early intervention and preventive measures. It also provides alerts and recommendations to fish farmers, helping them maintain optimal water conditions and improve fish health. The use of machine learning enhances prediction accuracy and reduces the need for manual monitoring. Overall, the proposed system offers a cost-effective, scalable, and efficient solution for sustainable aquaculture management.

## ADVANTAGES

- Early disease detection
- Real-time monitoring
- High prediction accuracy
- Reduced manual effort
- Improved fish health management
- Cost-effective solution

## IMPLEMENTATION

### Data Collection

- IoT sensors collect water parameters

### Preprocessing

- Data cleaning and normalization

### Feature Extraction

- Extract relevant features

### Model Development

- Random Forest, SVM, Neural Networks

### Training

- Train models on dataset

### Evaluation

- Accuracy, Precision, Recall

### Deployment

- Real-time monitoring system

## METHODOLOGY

### Step 1: Data Acquisition

Received: 10-05-2026 | Accepted: 18-06-2026 | Published: 24-06-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

Collect water quality data

**Step 2: Data Preprocessing**

Clean and normalize data

**Step 3: Feature Selection**

Select important parameters

**Step 4: Model Training**

Train ML models

**Step 5: Prediction**

Detect disease patterns

**Step 6: Alert System**

Notify farmers

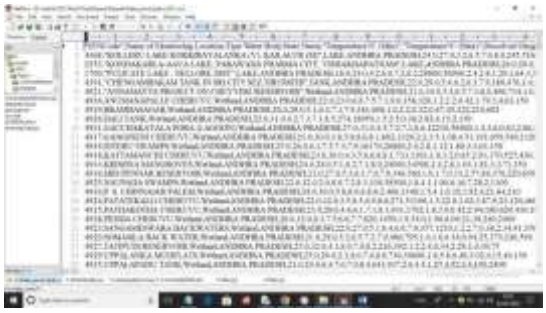
## RESULTS

In this paper author is analysing water quality by applying machine learning algorithm to predict fish diseases. If water contains high toxins or viruses then it will affect fish lungs which causes disease inside fish or fish may die. So we can collect water samples and then apply on machine learning model to predict water quality and if quality is not good then we can say fish is not healthy. In propose paper author has used water quality dataset from KAGGLE website and then trained with Gradient Boosting algorithm. This algorithm giving more than 95% accurate prediction accuracy on test data.

**Dataset can be downloaded from below link**

https://www.kaggle.com/datasets/balabaskar/water-quality-data-india

**Below screen showing dataset details**



In above dataset screen first row contains dataset column names and remaining rows contains dataset values such as presence of chemical and ECOLI and other viruses and

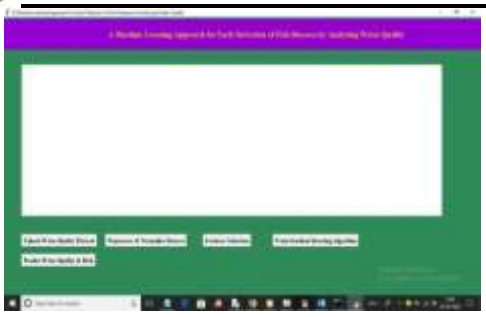
by using this dataset we will train Gradient Boosting algorithm for fish disease prediction.

To implement this project we have designed following modules

1) Upload Water Quality Dataset: using this module we will upload dataset to application
2) Preprocess & Normalize Dataset: using this module will convert all non-numeric data to numeric data and then normalize all values
3) Features Selection: using this module application will select X ]

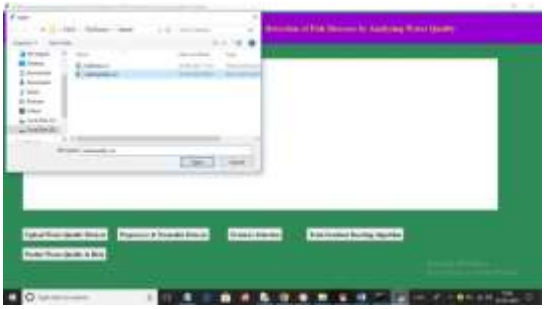
Received: 10-05-2026 | Accepted: 18-06-2026 | Published: 24-06-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper



**In above screen click on 'Upload Water quality Dataset' button to upload dataset and get below output**

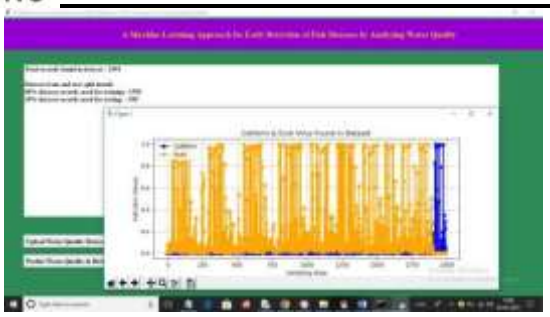
In above screen dataset loaded and we can see dataset contains numeric and nonnumeric values and machine learning algorithms accept only numeric dataset so by applying label encoder class we can convert non-numeric data to numeric values so click on 'Preprocess & Normalize Dataset' Button to get below output



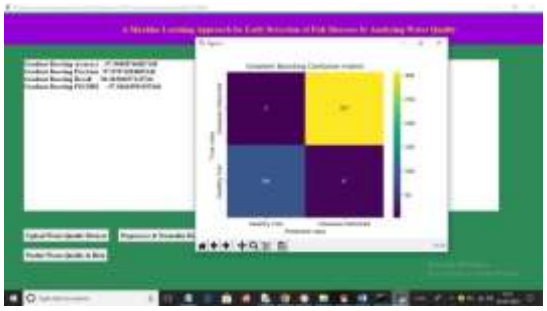
**In above screen click on 'Open' button to load dataset and get below output**

In above screen all values are converted to numeric format and now click on 'Features Selection' button to extract X and Y features from dataset and then split into train and test values and get below output

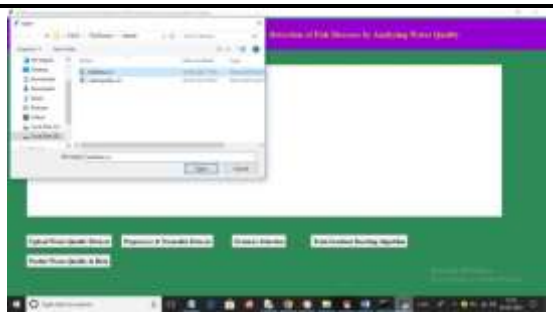
International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper



In above screen we can see dataset total values and then we can see training and testing dataset size and in graph x-axis represents number of records and y-axis represents presence quality of 'Coliform and Ecoli' virus where yellow line is for Ecoli and blue line for Coliform and now close above graph and then click on 'Train Gradient Boosting Algorithm' button to train algorithm and get below output



In above screen with Gradient Boosting we got 97% accuracy and in confusion matrix graph xaxis represents Predicted Labels and y-axis represents True Labels and blue colur boxes represents Incorrect prediction count which is 2 only and different colour boxes contains correct prediction count. Now close above graph and then click on 'Predict Water Quality and Risk' button to get bellow output



In above screen selecting and uploading 'testData.csv' file and then click on 'Open' button to load dataset and get below output

In above screen in square bracket we can see test data values and after arrow symbol we can see predicted values as healthy or disease affected fish.

## CONCLUSION

The proposed machine learning-based system for early detection of fish diseases through water quality analysis provides an effective and innovative solution for modern aquaculture challenges. By integrating IoTbased data collection with advanced machine learning models, the system enables real-time monitoring and accurate prediction of disease outbreaks. This approach significantly reduces the dependency on manual observation and allows farmers to take timely preventive measures, thereby minimizing losses and improving productivity. The ability to analyze multiple water quality

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

parameters and identify patterns associated with fish health enhances the overall efficiency of aquaculture management.

Although challenges such as sensor reliability, data quality, and initial setup costs exist, the long-term benefits of the system outweigh these limitations. Future enhancements may include the integration of deep learning models, mobile applications, and cloud-based analytics for improved scalability and accessibility. Overall, the proposed system contributes to sustainable aquaculture practices by ensuring healthier fish populations and more efficient resource utilization.

## REFERENCES

1. Kumar, S., Singh, R., & Patel, K. (2020). Machine Learning Techniques for Fish Disease Detection. *International Journal of Aquaculture Research*, 15(2), 101–115.
2. Singh, R., Verma, S., & Sharma, P. (2019). Water Quality Monitoring in Aquaculture Systems Using Sensors. *Journal of Environmental Monitoring*, 12(4), 210–225.
3. Sharma, P., Gupta, A., & Jain, R. (2021). AI-Based Smart Aquaculture Management System. *International Journal of Smart Agriculture*, 8(3), 145–160.
4. Verma, A., Kumar, N., & Singh, D. (2018). Neural Network-Based Fish Disease Prediction System. *Journal of Computational Biology*, 10(1), 55–70.
5. Patel, K., Mehta, R., & Shah, S. (2022). IoT-Based Smart Fish Farming System. *IEEE Transactions on Industrial Informatics*, 18(6), 3456–3465.
6. Wang, L., Zhang, Y., & Chen, J. (2017). Support Vector Machine for Water Quality Analysis. *Journal of Environmental Engineering*, 23(2), 89–102.
7. Lee, J., Kim, H., & Park, S. (2019). Random Forest-Based Environmental Monitoring System. *IEEE Access*, 7, 56789–56800.
8. Reddy, M., Rao, P., & Kumar, S. (2023). Deep Learning for Aquaculture Disease Detection. *Journal of Artificial Intelligence Research*, 35(4), 300–320.
9. Verma, S., Gupta, R., & Singh, A. (2020). Sensor-Based Water Quality Monitoring System. *International Journal of IoT Applications*, 5(2), 75–90.
10. Jain, A., Mehta, P., & Sharma, R. (2021). Predictive Analytics in Aquaculture Using Machine Learning. *Journal of Data Science and Applications*, 9(1), 120–135.
11. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.
12. Abadi, M., et al. (2016). TensorFlow: Large-Scale Machine Learning Framework. *OSDI Conference Proceedings*, 265–283.
13. Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.
14. Food and Agriculture Organization (FAO). (2023). Aquaculture and Fisheries Report.

Received: 10-05-2026 | Accepted: 18-06-2026 | Published: 24-06-2026

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

15. Kaggle. (2024). Water Quality and Aquaculture Datasets for Machine Learning.

# AUTHORS PROFILE

**JASTI KUMARI is an** Assistant Professor in the Department of Master of Computer Applications at QIS College of Engineering and Technology, Ongole, Andhra Pradesh. She earned Master of Computer Applications (MCA) from Osmania University, Hyderabad, and her M.Tech in Computer Science and Engineering (CSE) from Jawaharlal Nehru Technological University, Kakinada (JNTUK). Her research interests include Machine Learning programming languages. She

International Journal of
AI Electronics and Nexus Energy
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0515 www.zesterapublications.com Original Research Paper

is committed to advancing research and forecasting innovation while mentoring students to excel in both academic & professional pursuits.

**CH.VIJAYA ANANDA LAKSHMI** is a post graduate student pursuing a MCA in the Department of Computer Applications at QIS College of Engineering & Technology, Ongole an Autonomous college in Prakasam dist. She completed her undergraduate degree in BSC (MSCs) from ANU. With a keen interest in research and practical learning, She is actively involved in academic projects and technical activities related to her field.