JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Classification of water quality using shrimp pond by utilizing machine learning

A P Roger rozario [1] , K Veerakumar[2] , A Kathir vel[3] , J Salman[4] , A Santhosh[5]

[1]Assistant Professor, Department of EEE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamilnadu, India
[2,3,4,5] UG Students, Department of EEE, Sri Ramakrishna Institute of Technology, Coimbatore, Tamilnadu, India

***Abstract*:** Shrimp ponds, also known as temporary water bodies formed by shrimp during periods of high flow, have received limited attention in water quality assessment studies. In this study, we propose a classification framework that leverages machine learning algorithms to analyze water quality data collected from shrimp ponds. The framework includes feature selection, data preprocessing, and model training steps to effectively classify water quality based on key indicators such as pH, dissolved oxygen, turbidity, and nutrient concentrations. We demonstrate the effectiveness of our approach through experiments conducted on real-world water quality datasets collected from various shrimp ponds. Our results show promising performance in accurately classifying water quality levels, thus providing valuable insights for environmental monitoring and management efforts. The classification results demonstrate the effectiveness of the proposed approach in accurately predicting water quality levels in shrimp ponds. By integrating advanced data analytics techniques with domain knowledge in hydrology and ecology, this research contributes to the development of efficient strategies for monitoring and managing water resources in shrimp pond ecosystems.

1 INTRODUCTION

Water quality assessment is crucial for ensuring the health and sustainability of aquatic ecosystems. Shrimp ponds, which are artificial or semi-natural water bodies connected to streams or rivers, play a vital role in maintaining water quality and supporting diverse aquatic life. However, the classification of water quality in stream ponds presents unique challenges due to their dynamic nature, varying inputs, and interactions with surrounding landscapes.

In recent years, advancements in technology and the availability of sophisticated monitoring tools have enabled researchers and environmental practitioners to develop innovative approaches for classifying water quality in stream ponds. This technical paper aims to explore and discuss these approaches, focusing on the utilization of data-driven methods and remote sensing techniques for accurate and efficient classification.

The classification of water quality in shrimp ponds involves the assessment of various physicochemical parameters, such as dissolved oxygen, pH, temperature, turbidity, nutrient levels, and contaminants. Traditional methods rely on manual sampling and laboratory analysis, which can be time-consuming, costly, and limited in spatial and temporal coverage. In contrast, data-driven approaches leverage statistical modeling, machine learning algorithms, and sensor networks to process large volumes of data and extract meaningful insights in real-time.

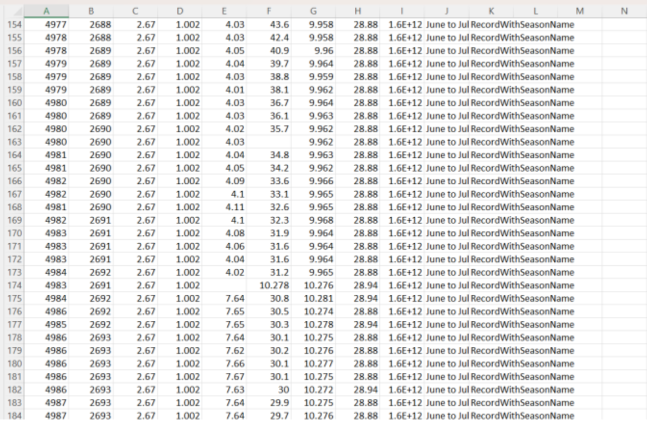
Remote sensing technologies, including satellite imagery, aerial drones, and ground-based sensors, offer valuable tools for monitoring water quality in shrimp ponds over large spatial scales. These techniques enable the detection of spatial patterns, temporal trends, and environmental changes that impact water quality dynamics. By integrating remote sensing data with in-situ measurements and hydrological models, researchers can develop robust classification models to assess and predict water quality conditions with high accuracy.

2. OBJECTIVE OF THE STUDY

The Present research study was carried out with the objective to study the water quality parameters of selected shrimp ponds. Information gathered from the study would be useful to understand the productivity of the shrimp culture ponds with reference to water characteristics during the summer crop (JUNE).

DATA COLLECTION

A Raw Dataset was collected from shrimp ponds device as a dataset.csv file [1], that contains more than 20 parameters and extract the features that are highly correlated to the labelled data.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
154	4977	2688	2.67	1.002	4.03	43.6	9.958	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
155	4978	2688	2.67	1.002	4.03	42.4	9.958	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
156	4978	2689	2.67	1.002	4.05	40.9	9.96	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
157	4979	2689	2.67	1.002	4.04	39.7	9.964	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
158	4979	2689	2.67	1.002	4.03	38.8	9.959	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
159	4979	2689	2.67	1.002	4.01	38.1	9.962	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
160	4980	2689	2.67	1.002	4.03	36.7	9.964	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
161	4980	2689	2.67	1.002	4.03	36.1	9.963	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
162	4980	2690	2.67	1.002	4.02	35.7	9.962	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
163	4980	2690	2.67	1.002	4.03		9.962	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
164	4981	2690	2.67	1.002	4.04	34.8	9.963	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
165	4981	2690	2.67	1.002	4.05	34.2	9.962	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
166	4982	2690	2.67	1.002	4.09	33.6	9.966	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
167	4982	2690	2.67	1.002	4.1	33.1	9.965	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
168	4981	2690	2.67	1.002	4.11	32.6	9.965	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
169	4982	2691	2.67	1.002	4.1	32.3	9.968	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
170	4983	2691	2.67	1.002	4.08	31.9	9.964	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
171	4983	2691	2.67	1.002	4.06	31.6	9.964	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
172	4983	2691	2.67	1.002	4.04	31.6	9.964	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
173	4984	2692	2.67	1.002	4.02	31.2	9.965	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
174	4983	2691	2.67	1.002		10.278	10.276	28.94	1.6E+12	June to Jul	RecordWithSeasonName			
175	4984	2692	2.67	1.002	7.64	30.8	10.281	28.94	1.6E+12	June to Jul	RecordWithSeasonName			
176	4986	2692	2.67	1.002	7.65	30.5	10.274	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
177	4985	2692	2.67	1.002	7.65	30.3	10.278	28.94	1.6E+12	June to Jul	RecordWithSeasonName			
178	4986	2693	2.67	1.002	7.64	30.1	10.275	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
179	4986	2693	2.67	1.002	7.62	30.2	10.276	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
180	4986	2693	2.67	1.002	7.66	30.1	10.277	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
181	4986	2693	2.67	1.002	7.67	30.1	10.275	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
182	4986	2693	2.67	1.002	7.63	30	10.272	28.94	1.6E+12	June to Jul	RecordWithSeasonName			
183	4987	2693	2.67	1.002	7.64	29.9	10.275	28.88	1.6E+12	June to Jul	RecordWithSeasonName			
184	4987	2693	2.67	1.002	7.64	29.7	10.276	28.88	1.6E+12	June to Jul	RecordWithSeasonName			

Figure 1: Dataset Collected from Shrimp POND.

DATASET DESCRIPTION

A comprehensive dataset focusing on water quality in stream ponds encompasses a myriad of parameters crucial for understanding ecosystem health and human impact. This dataset typically includes physical, chemical, and biological indicators such as temperature, pH, dissolved oxygen, conductivity, specific gravity, total dissolved solid. Each parameter provides insights into various aspects of water quality, reflecting natural processes, anthropogenic influences, and potential ecological risks. By analyzing trends and fluctuations in these indicators over time and across different locations, researchers can assess the overall condition of stream ponds, identify potential sources of contamination, evaluate the effectiveness of management strategies, and support informed decision-making for conservation and water resource management efforts. Such a dataset serves as a valuable resource for scientists, policymakers, and environmental professionals working towards the preservation and restoration of freshwater ecosystems.

3. MATERIALS AND METHODS

3.1 Method of water samples collection

The water sample was collected between 6-7 am in all selected culture ponds by dipping 500ml clean polythene bottles 1-2 feet depth in the ponds and samples were brought to the laboratory for analysis of various chemical parameters like Salinity, pH, specific gravity, Total Dissolved Solids (TDS), dissolved oxygen by standard methods according to APHA.

3.2 Temperature

Water temperature was measured with a mercury-filled Celsius thermometer ranging 0 to 50 °C. To measure temperature the thermometer was dipped in to the water for one minute and the stable temperature final reading was recorded.

3.3 pH

Water pH of the collected samples was measured using a digital pH meter (edge Bluetooth Smart pH Electrode and Meter - HI2202 HANNA instruments) nearest to 0.01. Before using the instrument it was calibrated with pH 7 and pH 10 buffer solutions. The pH probe was immersed in the water samples to be tested without exceeding the maximum immersion level. Then the sample was stirred gently and waited for the reading to stabilize and the final pH reading was recorded.

3.4 Dissolved Oxygen (DO)

Dissolved Oxygen plays an important role in growth and production through its direct effect on feed consumption and maturation. Dissolved oxygen affects the solubility and availability of many nutrients in the pond water. Low level of dissolved Oxygen can cause damages in the oxidation state of substances from the oxidized to the reduced form lack of dissolved oxygen can be directly harmful to shrimps and cause a considerable increase within the level of hepatotoxic metabolic performances in shrimp and can cut back growth and molting and cause stress its leads to mortality.

3.5 Salinity

Salinity of the collected water samples was measured using a digital refractometer model Seawater Analysis HI96822 Hanna instruments. The salinity probe was immersed in the water samples to be tested without exceeding the maximum immersion level and waited for the salinity reading, and the final salinity value was recorded.

3.6 Specific gravity

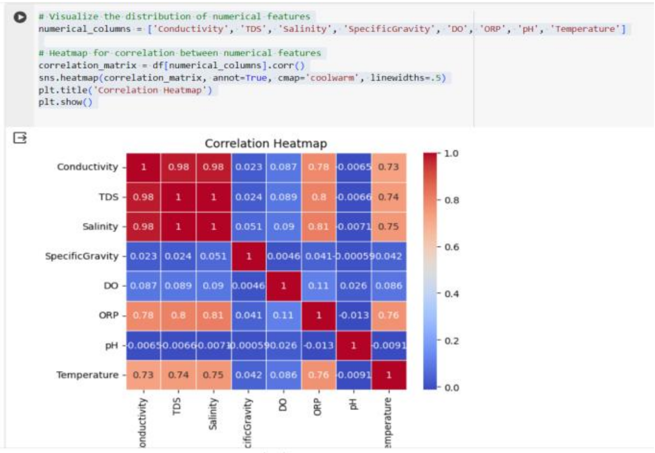
The ratio of the density of a substance to the density of a reference substance (usually water), reflects the concentration of dissolved solids, suspended particles, and other impurities in water. Changes in specific gravity can indicate variations in water composition, salinity, and contaminant levels, which are critical factors affecting water quality in stream ponds. By incorporating specific gravity measurements into classification models, researchers can improve the detection and characterization of environmental stressors, pollution sources, and ecosystem health indicators.

3.7 Total Dissolved Solids (TDS)

It represent the concentration of inorganic and organic substances dissolved in water and serve as a key parameter for evaluating water quality. The measurement and analysis of TDS levels provide valuable insights into the chemical composition, nutrient dynamics, and pollution sources affecting stream pond ecosystems.

This technical paper aims to review the significance of TDS parameters for water quality assessment in stream ponds, focusing on measurement techniques, interpretation, and practical applications. By understanding TDS levels, researchers can effectively monitor and manage water resources, ensuring the health and sustainability of stream pond ecosystems.

]



```
# Visualize the distribution of numerical features
numerical_columns = ['Conductivity', 'TDS', 'Salinity', 'SpecificGravity', 'DO', 'ORP', 'pH', 'Temperature']

# Heatmap for correlation between numerical features
correlation_matrix = df[numerical_columns].corr()
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', linewidths=.5)
plt.title('Correlation Heatmap')
plt.show()
```

Correlation Heatmap

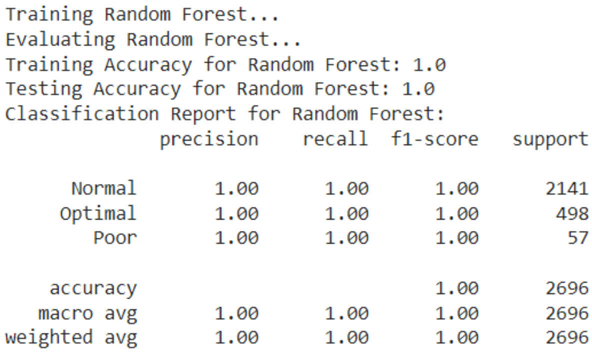
	onductivity	TDS	Salinity	cificGravity	DO	ORP	pH	emperature
Conductivity	1	0.98	0.98	0.023	0.087	0.78	-0.0065	0.73
TDS	0.98	1	1	0.024	0.089	0.8	-0.0066	0.74
Salinity	0.98	1	1	0.051	0.09	0.81	-0.0071	0.75
SpecificGravity	0.023	0.024	0.051	1	0.0046	0.041	-0.00059	0.042
DO	0.087	0.089	0.09	0.0046	1	0.11	0.026	0.086
ORP	0.78	0.8	0.81	0.041	0.11	1	-0.013	0.76
pH	-0.0065	-0.0066	-0.0071	-0.00059	0.026	-0.013	1	-0.0091
Temperature	0.73	0.74	0.75	0.042	0.086	0.76	-0.0091	1

Fig 2: Correlation heatmap

4.METHODOLOGY

4.1 Random Forest:

Random forest (RF) is a relatively new model, developed by Leo Breiman and Adele Cutler in the end of 90s. At each split of the observed sample data, a random subset of variables is selected and the process is repeated until the specified number of decision trees is generated. Each tree is built from a bootstrap sample drawn with replacement from the observed data, and the predictions of all trees are finally aggregated through majority voting. A feature of RFs is the definition of an out-of-bag (OOB) error, which is calculated from observations that were not used to build a particular tree; it can thus be considered as an internal cross-validation error measure. This is an important feature for the type of experiments carried out in this study, because it simplifies the otherwise cumbersome cross-validation procedures that would be required if alternative classification methods such as, for instance, support vector machines or artificial neural networks were used.

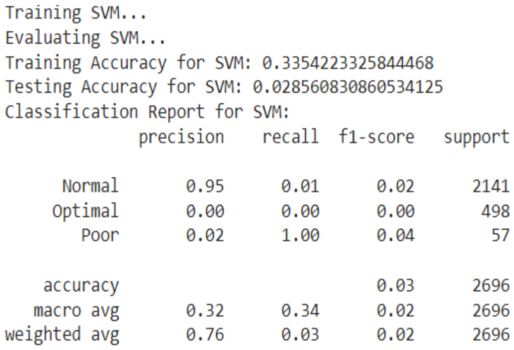


```
Training Random Forest...
Evaluating Random Forest...
Training Accuracy for Random Forest: 1.0
Testing Accuracy for Random Forest: 1.0
Classification Report for Random Forest:
              precision    recall  f1-score   support

      Normal       1.00      1.00      1.00      2141
     Optimal       1.00      1.00      1.00       498
        Poor       1.00      1.00      1.00        57

    accuracy                           1.00      2696
   macro avg       1.00      1.00      1.00      2696
weighted avg       1.00      1.00      1.00      2696
```

Fig 3: Performance metrics of Model using Random forest Algorithm

4.2 Support vector machine (svm):

The Support Vector Machine (SVM) stands as a potent and flexible supervised machine learning algorithm. Renowned for its proficiency in both classification and regression assignments, and aims to discover an optimal hyperplane that effectively divides data into separate classes, ensuring a distinct margin between them. Its strength lies in tackling intricate decision boundaries and handling high-dimensional data effectively.SVM functions by recognizing support vectors, which represent the data points nearest to the decision boundary. These support vectors hold significant importance in delineating the separation margin and enhancing the classifier's effectiveness. The algorithm aims to maximize this margin, guaranteeing that the decision boundary maintains a maximum distance from the support vectors.

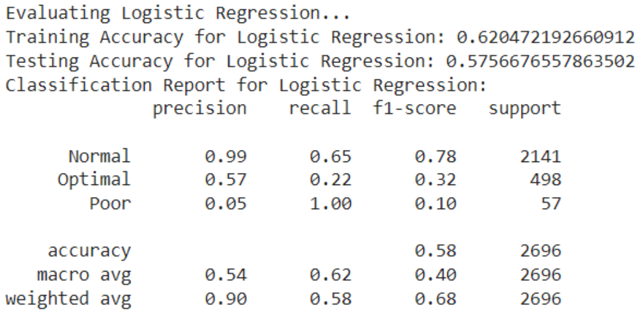


```
Training SVM...
Evaluating SVM...
Training Accuracy for SVM: 0.3354223325844468
Testing Accuracy for SVM: 0.028560830860534125
Classification Report for SVM:
              precision    recall  f1-score   support

      Normal       0.95      0.01      0.02      2141
     Optimal       0.00      0.00      0.00       498
        Poor       0.02      1.00      0.04        57

    accuracy                           0.03      2696
   macro avg       0.32      0.34      0.02      2696
weighted avg       0.76      0.03      0.02      2696
```

Fig 4: Performance metrics of Model using SVM Algorithm

4.3 Logistic regression algorithm:

Logistic Regression stands as a fundamental and extensively utilized machine learning algorithm, primarily employed in binary classification tasks. Contrary to its name, it functions as a classification algorithm rather than a regression technique. It predicts the probability of a binary result (1 or 0) by incorporating one or more predictor variables. Operating as a linear algorithm, it utilizes the logistic function (sigmoid function) to model these probabilities. The algorithm functions by computing the weighted sum of input features and then using the logistic function to generate a probability value ranging from 0 to1. This probability signifies the chance of the event happening (class 1). Logistic Regression is widely recognized for its simplicity, interpretability, and efficiency. It is also extensible to multi-class classification problems through techniques like one-vs-rest (OvR). In binary classification, a threshold is applied to the predicted probabilities to assign data points to one of the two classes. The choice of threshold impacts the trade-off between precision and recall. Logistic Regression is commonly used in fields where binary classification is essential, such as medical diagnosis, spam detection, and credit scoring.



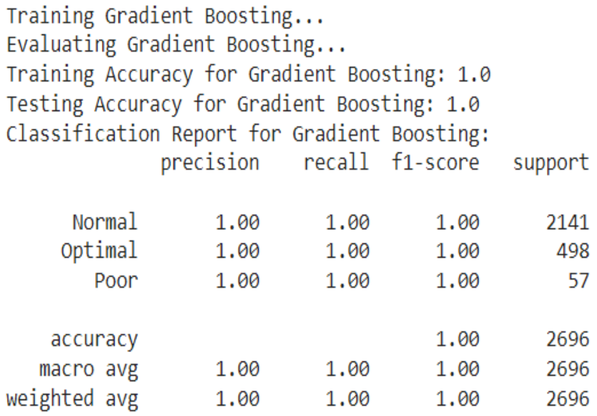
```
Evaluating Logistic Regression...
Training Accuracy for Logistic Regression: 0.620472192660912
Testing Accuracy for Logistic Regression: 0.5756676557863502
Classification Report for Logistic Regression:
              precision    recall  f1-score   support

      Normal       0.99      0.65      0.78      2141
     Optimal       0.57      0.22      0.32       498
        Poor       0.05      1.00      0.10        57

    accuracy                           0.58      2696
   macro avg       0.54      0.62      0.40      2696
weighted avg       0.90      0.58      0.68      2696
```

Fig 5: Performance metrics of Model using Logistic regression algorithm .

4.4 Gradient boosting algorithm:

The gradient boosting algorithm is a machine learning technique used for regression and classification tasks. It works by combining multiple weak learners, often decision trees, to create a strong predictive model. In gradient boosting, each new model is trained to correct the errors made by the ensemble of previous models. The algorithm iteratively fits new models to the residuals of the ensemble, gradually reducing the overall error. This process continues until a specified number of models have been added or until a certain level of performance is achieved. At its core, gradient boosting involves optimizing a loss function by iteratively adding new models to the ensemble. In each iteration, the algorithm fits a new tree to the residual errors of the current ensemble. This process continues until a specified number of trees is reached or until a certain level of performance is achieved.

One popular variant of gradient boosting is the Gradient Boosting Machine (GBM), which was introduced by Jerome Friedman. GBM uses gradient descent optimization to minimize the loss function, updating the ensemble's predictions with each new tree added. The key idea behind gradient boosting is to combine the predictions of multiple weak learners to create a strong learner that generalizes well to unseen data. By iteratively refining the model's predictions, gradient boosting can often achieve high levels of accuracy and generalization performance.

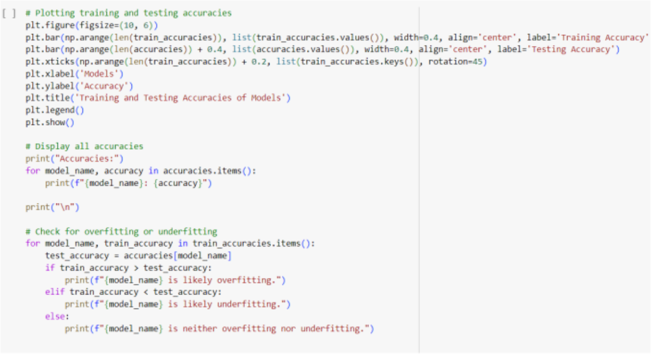


```
Training Gradient Boosting...
Evaluating Gradient Boosting...
Training Accuracy for Gradient Boosting: 1.0
Testing Accuracy for Gradient Boosting: 1.0
Classification Report for Gradient Boosting:
              precision    recall  f1-score   support

      Normal       1.00      1.00      1.00      2141
     Optimal       1.00      1.00      1.00       498
        Poor       1.00      1.00      1.00        57

    accuracy                           1.00      2696
   macro avg       1.00      1.00      1.00      2696
weighted avg       1.00      1.00      1.00      2696
```

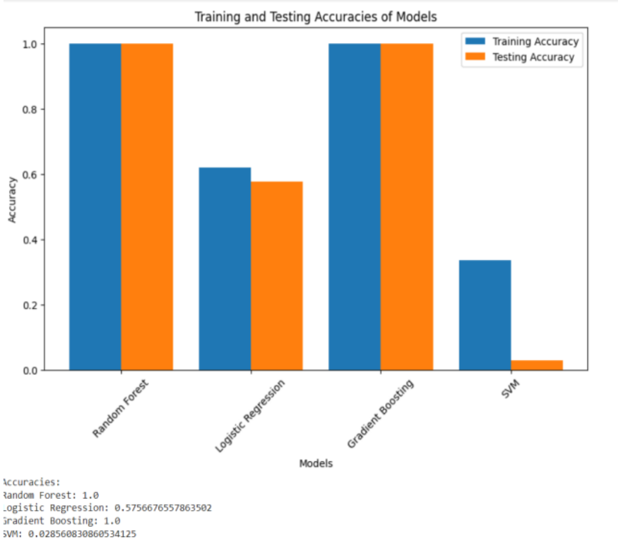
Fig 6: Performance metrics of Model using Gradient boosting algorithm

5. RESULT AND ANALYSIS

The analysis of physical parameters such as temperature, conductivity, a indicated fluctuations influenced by seasonal variations, land use patterns, and hydrological conditions. These variations can impact the thermal regime of the water body, its ability to support aquatic life, and the clarity of the water, which are crucial factors for overall ecosystem health. Chemical parameters, including pH, dissolved oxygen, nutrients (nitrogen and phosphorus), and heavy metals, exhibited spatial and temporal variability, reflecting natural processes such as nutrient cycling, as well as anthropogenic inputs from agricultural runoff, urban discharge, and industrial activities. the presence of organic pollutants, such as pesticides, pharmaceuticals, and industrial contaminants, highlighted potential sources of pollution and raised concerns about water quality degradation and human health risks. The dataset also included microbial content measurements, revealing the presence of fecal coliforms and other pathogens, indicating potential contamination from sewage discharge or agricultural runoff, which can compromise water safety and pose health hazards for recreational activities and drinking water sources. Overall, the detailed analysis of these parameters provided a comprehensive understanding of the current state of water quality in stream ponds, identifying both natural variability and anthropogenic impacts.



```
[ ] # Plotting training and testing accuracies
    plt.figure(figsize=(10, 6))
    plt.bar(np.arange(len(train_accuracies)), list(train_accuracies.values()), width=0.4, align='center', label='Training Accuracy'
    plt.bar(np.arange(len(accuracies)) + 0.4, list(accuracies.values()), width=0.4, align='center', label='Testing Accuracy')
    plt.xticks(np.arange(len(train_accuracies)) + 0.2, list(train_accuracies.keys()), rotation=45)
    plt.xlabel('Models')
    plt.ylabel('Accuracy')
    plt.title('Training and Testing Accuracies of Models')
    plt.legend()
    plt.show()

    # Display all accuracies
    print("Accuracies:")
    for model_name, accuracy in accuracies.items():
        print(f"{model_name}: {accuracy}")

    print("\n")

    # Check for overfitting or underfitting
    for model_name, train_accuracy in train_accuracies.items():
        test_accuracy = accuracies[model_name]
        if train_accuracy > test_accuracy:
            print(f"{model_name} is likely overfitting.")
        elif train_accuracy < test_accuracy:
            print(f"{model_name} is likely underfitting.")
        else:
            print(f"{model_name} is neither overfitting nor underfitting.")
```



```
Accuracies:
Random Forest: 1.0
Logistic Regression: 0.5756676557863502
Gradient Boosting: 1.0
SVM: 0.028560830860534125
```

Fig 8: Comparison of accuracies of all algorithm

6. CONCLUSION

The assessment of water quality using shrimp ponds as indicators offers valuable insights into the health and integrity of freshwater ecosystems. The comprehensive analysis of physical, chemical, and biological parameters reveals a complex interplay of natural processes and human activities influencing water quality dynamics. The results highlight the vulnerability of shrimp ponds to various stressors, including nutrient pollution, organic contaminants, and microbial pathogens, which can compromise ecosystem functioning and pose risks to human health Integrated approaches that incorporate land-use planning, pollution control measures, and ecosystem restoration efforts are essential for mitigating anthropogenic impacts and promoting sustainable water management practices. The conservation and restoration of shrimp ponds are essential for maintaining biodiversity, supporting ecosystem services, and ensuring the availability of clean water for both ecological and human needs. By prioritizing the protection of these valuable freshwater habitats and implementing science-based management approaches, we can work towards enhancing water quality, preserving ecosystem health, and fostering resilience in the face of environmental challenges.

7. FUTURE WORKS

In the future, the proposed system could be expanded and enhanced in several keys areas to advance our understanding and management of freshwater ecosystems.

A. *Long-term Monitoring:*

Establishing and maintaining long-term monitoring programs to track trends and changes in water quality parameters over extended periods. This would provide valuable data for assessing the efficacy of management interventions, identifying emerging threats, and understanding the impacts of climate change on stream pond ecosystems.

B. *Ecological Responses to Stressors:*

Studying the ecological responses of stream ponds to multiple stressors, including pollution, habitat alteration, invasive species, and climate change. Integrating ecological assessments with water quality monitoring can provide insights into ecosystem resilience and inform adaptive management strategies.

C. *Spatial Modeling and Prediction:*

Developing spatially explicit models and predictive tools to assess and forecast water quality dynamics in shrimp ponds. Integrating environmental data with modeling techniques such as machine learning, remote sensing, and spatial statistics can enhance our ability to identify drivers of water quality variation and anticipate future changes.

D. *Emerging Contaminants:*

Investigating the occurrence, fate, and effects of emerging contaminants such as pharmaceuticals, personal care products, and microplastics in shrimp ponds. Understanding the potential risks posed by these contaminants to aquatic organisms and ecosystem health is crucial for developing targeted mitigation strategies.

REFERANCE

[1] APHA, (American Public Health Association). Standard methods for the examination of water and waste water, New York, 2005.

[2] Boyd CE, AW Fast. Pond monitoring and management. In: Marine Shrimp Culture: Principles and Practices (eds. A.W. Fast and L. James Lester). Elsevier Science Publishers BV, 2009.

[3] Boyd CE. Water quality for pond aquaculture. Research and Development Series No. 43. International Center for Aquaculture and Aquatic Environments, Alabama Agricultural Experimental Station, Auburn University, Auburn, Alabama, 20218.

[4] Boyd CE, BW Green. Coastal water quality monitoring in shrimp farming areas, an example from Honduras. Report prepared under the World Bank, NACA, WWF and FAO Consortium Program on Shrimp Farming and the Environment, published by the Consortium, 2017.

[5] APHA (American Public Health Association), Standard methods for the examination of water and waste water 15th edition, American Public Health Association, Washington D.C., USA, 1998,

[6] Wyban JA, Ogle J, Pruder GD, Rowland LW, Leung PS. Design, operation, and comparative financial analysis of shrimp farms in Haweii and Taxes. Tech Report 86-6. The Oceanic Institute, Honolulu, Haweii, USA, 2009.

[8] Parry G. Excetion. in T.H. Waterman, editor, the physiology of crustacean, Academic Press, New York. 2016.

[9] Bower CE, Bidwell JP. Jonization of ammonia in seawater, effects of temperature, pH and salinity. Journal of Fisheries Research Board Canada. 2016.