International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

AI-Based Principal Component Analysis (PCA) Approach for the Determination of Key Water Quality Parameters

Dinesh M.*[1], Prakash S.[2], Anasuya N. Jadagerimath [3]

Submitted: 24/04/2023 **Revised:** 28/06/2023 **Accepted:** 06/07/2023

Abstract: The health of the fish is impacted by several water quality indicators, including pH, Dissolved Oxygen, Unionized Ammonia, Turbidity, etc. Therefore, research has been conducted utilizing a machine learning approach to determine the key variables that have the greatest impact on fish. To recognize this, a new classification system called Fish Classification Index (FCI) is developed to classify fish environments. In this approach, the water body ecology is divided into four groups: Alive, Disaster, Just to Dwell (JTD), and Just to Maintain (JTM). A mathematical model is developed based on this Fish Classification Index. Around ten water quality criteria are considered for the calculation in this paper. Additionally, it's critical to comprehend which factor has the greatest impact on fish health and the degree to which it does so. Therefore, Principal Component Analysis (PCA), a statistical method, is utilized in Artificial Intelligence (AI) to investigate the key variables that affect the Fish Classification Index's judgment. Using this technique, it was observed that pH and DO each have a 22% and 32% influence on the decision, respectively. Overall, it has been discovered that four parameters account for about 80% of decision-making, with the remaining six parameters present in water accounting for 20% of decision-making. The Identification and comprehension of the most important factors from the current research will help the relevant stakeholders to act appropriately in maintaining a healthy water body ecology. Additionally, it aids engineering designers in making appropriate use of the right sensors, leading to efficient resource use.

***Keywords*:** *Artificial Intelligence; Machine Learning; Water quality parameters; Aqua farming; Aquaculture*

1. Introduction

Fishes are raised in enclosures for food production in a practice known as fish farming. In terms of producing animal food, it is expanding the quickest. Today, these synthetic settings are used to raise roughly half of the fish consumed worldwide. [1] Fish culture is a productive approach for producing protein-rich foods from aquatic environments and is a rich source of animal protein. The primary function of fish farming is to raise the nutritional standards of the populace. Utilizing resources like water and land is made easier by fish culture. It encourages the development of other ancillary industries in the nation. Aquatic resources (marine and inland) are abundant and diverse in India, making them suitable for aquaculture and capture fisheries. Inland water bodies are extensively used for both culture and capture fisheries, whereas marine water bodies are primarily exploited for those resources.

The majority of inland water bodies have captive ecosystems [2] that allow for significant human involvement in biological production, which has the potential to greatly boost fish productivity. Freshwater bodies such as rivers, canals, streams, lakes, floodplain wetlands or beels (ox-bow lakes, back swamps, etc.), reservoirs, ponds, tanks, and other abandoned water bodies are considered to be part of the inland water system. All of the fish's physiological functions including breathing, waste excretion, eating, preserving the equilibrium of the salts, and reproduction are controlled by the quality of water. Water quality so determines whether an aquaculture operation is successful or unsuccessful. The ongoing deterioration [3] of Water Resources from anthropogenic sources requires criteria for choosing aquaculture sites that use water quality as a basis. Important water quality parameters that need to be taken care of are Dissolved Oxygen (DO), pH, Turbidity, Unionized Ammonia, BOD of water, Bacteria Count, Amount of Phosphate, Nitrite and Nitrate content of water, alkalinity of water. For optimal fish growth, a given water body must have a Real-Time Water Quality Monitoring System. [4] The amount of dissolved oxygen accessible to fish is influenced by a number of factors. [5] Dissolved oxygen levels are also impacted by microorganism populations. Phytoplankton is one type of microbe that can carry out photosynthesis. Light from the sun and carbon dioxide are changed into oxygen and organic material during photosynthesis. The amount of dissolved oxygen in the water is increased by this process. Respiration is another process carried out by phytoplankton and other microbes that use oxygen and depletes the oxygen supply. Depending on the metabolic functions that different populations of microorganisms carry out, the oxygen

1Ph.D. Scholar, Visvesvaraya Technological University, Belagavi-590018, Karnataka, India
ORCID ID : 0000-0002-4345-434X
2Professor of Computer Science & Engineering, East Point Group of Institutions, Bangalore- 560049, Karnataka, India
ORCID ID : 0000-0001-7147-4375
3Professor and HoD CSE(AI&ML), Don Bosco Institute of Technology Bengaluru-560074, Karnataka, India
ORCID ID : 0000-0002-8623-5948
** Corresponding Author Email: dinesh.svdm@gmail.com*

availability will change. Raising a lot of fish in a little amount of water is called aquaculture. Due to this artificially induced overpopulation, resources, such as dissolved oxygen, are quickly depleted. For this reason, artificial aeration and other techniques are used in fish farming operations to supplement the natural oxygen supply. Using dissolved oxygen sensors to monitor dissolved oxygen in aquafarming applications has several benefits. Monitoring allows managers to be notified when dissolved oxygen drops too low, endangering the health of fish or shellfish. Monitoring also provides feedback to control systems, which can automatically regulate oxygen in aquaculture via mechanical aeration.[6] Careful control of aeration systems is critical because these systems require significant energy expenditure. Optimizing aeration for sustainable growth conditions saves energy and saves money.

Fish survival, growth, and reproduction can be affected by pH levels.[7] pH extremes have the potential to kill fish. The pH may also have an impact on populations of microorganisms that affect the quality of the water. Some microbes, like phytoplankton, carry out crucial metabolic processes like photosynthesis. Photosynthesis is a significant source of dissolved oxygen, as was previously mentioned. Aquaculture habitats can be kept at the right pH to allow helpful bacteria to outcompete harmful ones.[8] This improves the ecosystem's circumstances for fish growth, reproduction, and survival. The pH can change for a variety of reasons in fish and shellfish habitats. Acid rain and agricultural run-off both have the potential to harm water. pH can also be impacted by carbon dioxide emissions since it lowers pH when it dissolves in water and forms carbonic acid.[9] The use of pH sensors and transmitters for pH monitoring aids in determining when modifications are necessary.

From the above statements it can be inferred that water quality parameters are very crucial for the growth and well-being of the fish. As it’s known that machine learning is used in several problems similarly it can be used in aqua farming. A machine learning approach can be used in this scenario to classify various parameters for their appropriate relevancy. [10] Using training data with labels, supervised learning builds supervised models that can be used to forecast results for upcoming datasets. Unsupervised learning is a sort of machine learning task in which there are no labels or categories applied to the training data.[11] Machine learning techniques like supervised learning require labelled training data. For instance, the training data might consist of photos of handwritten digits combined with the labels (i.e., the right digit for each image) if we were training a supervised learning system to categorize handwritten digits. Assisted learning algorithms create a model from the training data that they can use to predict outcomes from new data. Results for upcoming datasets that are comparable to the labelled datasets can be predicted using the supervised machine learning job. Labels are used in supervised models as a guide to assist build an accurate model.[12] Examples of supervised learning issues are shown below. Supervised Machine Learning Program diagrammatically can be depicted as shown in the below figure 1

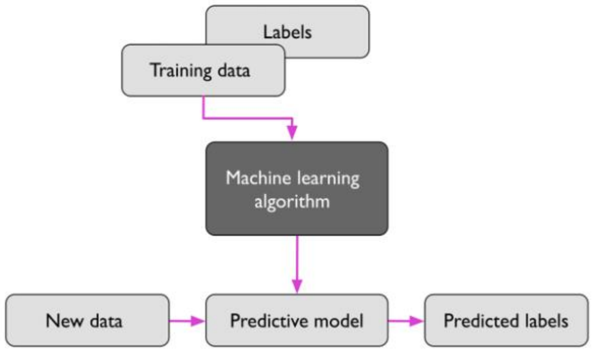

Fig.1. Supervised Machine Learning

Unsupervised learning, a subset of machine learning, is concerned with enabling a computer to make inferences about data on its own without the use of labels or classifications.[13][14] Contrastingly, supervised learning includes giving computer training data that has already been labelled Principal component analysis (PCA) and k-means clustering are two popular unsupervised learning algorithms. Numerous tasks, including dimensionality reduction, market segmentation, and novelty detection, can be accomplished via unsupervised learning. For business owners who wish to better understand their consumers and base their decisions on this knowledge,[16] the unsupervised machine learning job is frequently used to find hidden patterns or correlations within a dataset.

In this paper, the fish survival conditions in a given water body ecosystem are evaluated. It is not known how and to what extent each of the water quality parameters impacts the living conditions of the fish in the water. To assess this, a novel approach was devised to study the parameters, and an AI-based Principal Component Analysis (PCA)[15] technique is explored by using machine learning programming. This helps in identifying the crucial parameters that impact the life of the fish in water. The result will help engineers working on IoT development to choose the right number of sensors to use and also to decide on what sensors to be used. This will also assist the lake development authority in deploying the key parameters to be used with optimism.

The remaining section of the paper is organized as shown below. A review of the literature is included in Section 2 to determine whether any categorization techniques have been used for measurements of the water quality and, subsequently, the survival conditions of the fish. Additionally, it covers the use of machine learning in aquaculture. More information on the new fish classification system and its intended usage is provided in Sections 3 through 4. The graph and its association are examined in Section 5 before the conclusion's findings and suggestions for additional research are presented in Section 6.

2. Literature Survey

Aquafarming is experiencing a lot of innovative technological advancements. In addition to embedded systems, IoT and AI are quite clearly entering this field. This still has a long way to go. Table 1 below provides a literature review of how artificial intelligence (AI) has been integrated into water quality management systems.

Table 1. Literature Survey

Sl No	**Research Paper Reference**	**Year of Publication**	**Experiment and Results**	**Remarks**
1	[12]	2016	The purpose of this study was to create a water quality prediction model utilising Artificial Neural Networks (ANN) and time-series analysis to incorporate water quality parameters. They have used only 4 water quality parameter in this analysis.	Only 4 Water Quality Parameters are studied in the paper. From this paper, it can be inferred that there can be a more user-friendly approach in calculating the effectiveness of water quality.
2	[18]	2019	The findings of multiple linear regression (MLR), artificial neural network (ANN), and support vector machines (SVM) data mining methods were compared to those obtained using the PCA-based response surface regression (RSR) approach for calculating surface water quality parameters.	The water quality model is specific to the relevant region and specific season. It's the limitation observed.
3	[19]	2011	Here the water quality parameters are measured for a mangrove region. Around 14 water quality parameters were taken into account.	Here it has been tried to measure the water quality in the nearby distance of sea and land. The data is also evaluated based on the sea vicinity i.e. the distance of freshwater resources from the sea.
4	**Error! Reference source not found.**	2020	The paper's goal is to evaluate how the Hlaing River's water quality varies during the rainy and dry seasons. A range of physiochemical water quality indicators, such as Total Suspended Solid (TSS), Chemical Oxygen Demand (COD), and Biochemical Oxygen Demand (BOD), are measured. To ascertain the correlations between the variables and how many significant components are present in the data, principal	The analysis seems to be quite complex mathematical calculations. Not so user-friendly for the common man.

			component analysis is employed in SPSS.	
5	[21]	2023	Here the author has presented the Smart IOT-based water quality assessment. Also, they have used the PCA tool to assess the water quality.	The objective seems to be identifying the water quality for drinking and irrigation purposes. It uses the PCA technique.
6	[22]	2021	This research focuses on assessing irrigation water quality and creating a model for Irrigation Water Quality Index (IWQI) prediction based on salinity and sodicity. The parameters like Na^+ , Cl^- , EC, HCO^{3-} and SAR were measured.	Here they have used IoT-based classification for agriculture to assess the salinity to prevent damage to crops. Not much was done on the Aquafarming side.
7	[23]	2020	The paper speaks about remote sensing methods for evaluating the water quality.Parameters like Turbidity, Total Suspended sediments, Total Phosphorus and Total Nitrogen contents in the water is determined.	Basic water quality parameters are assessed. A new method has been advised for future requirements.
8	[24]	2018	To increase the effectiveness and safety of WRRFs, the author devised a flexible and reliable monitoring soft sensor technique to detect and identify anomalous influent data. The principal component analysis (PCA) and k-nearest neighbor (KNN) schemes' beneficial properties are combined in the suggested data-driven soft sensor technique. While KNN distances showed greater detection capability, robustness to underlying data distribution, and efficiency in processing high-dimensional datasets, PCA conducted efficient dimension reduction and identified correlations between inflow measurements.	The system proposed measures the water quality measurements at the water resource recovery facility. It provides another approach to water quality measurements. Not much is applicable to Aqua farming as such.

The literature review makes it clear that the majority of developments and efforts are put toward tracking the water quality metrics.[12]; in this instance, the author used and recommended only 4 water-related characteristics and employed the ANN approach for prediction in their research. In [22], the author utilizes an IoT-based approach to assess water quality for irrigation purposes. In [24], the author demonstrates how PCA and KNN algorithms were employed to evaluate the water monitoring system. From all of these factors, it is clear that water quality management systems are in use, but neither an accurate evaluation of the water quality of any ecosystem inside a water body nor a forecast of the future circumstances based on that water quality is performed for aquaculture.

3. Design and Methodology

The number of phytoplankton and zooplankton in the water, as well as environmental factors like pressure and temperature, were all carefully examined. The impact of the elements was assessed on basic water quality evaluations. For our reference study, we chose locally produced fish species including Catla, Rohu, Mrigal, etc. The minimum age of the fish considered into account was greater than one month old, or between three and five centimeter in length when growth was taken into account. Few indicators for the water quality management system were temperature, dissolved oxygen, water pH, Turbidity, Bio-Chemical Oxygen Demand (BOD), Nitrate, Nitrite, and the concentration of unionized ammonia (UIA) in the water. To choose the right living arrangements for the fish, we must first evaluate the optimal cozy settings for them. Once again, research was done to understand the various habitats that the stated fish species inhabit. The ideal parameters indicated in Table 2 were determined by research and evaluation of fish life conditions in various lakes.

Table 2. Suitable Values of Parameters

Sl No	Parameter	Preferred Interval	Not Comfortable / Distress Interval
1	Temperature (°C)	20-30	30 < Temp > 12
2	Dissolved Oxygen(mg/l)	5	8 < DO > 5
3	pH	6.5 – 9.0	11 < pH >4
4	UIA(mg/l)	0 – 0.025	Ammonia > 0.3
5	Turbidity (cm)	30-80	<12 or >80
6	Phosphate (mg/l)	0.03 – 2	>3
7	BOD(mg/l)	3-6	>10
8	Nitrate (mg/l)	0.1 – 4.5	>100 and <0.01
9	Nitrite (mg/l)	<0.02	>0.2
10	Alkalinity (mg/l)	25-100	<20 or >300

The optimum value was divided into 4 categories in order to improve the clarity and predictability of the outcomes.

1. **ALIVE**: This means that the fish may live peacefully and that the ecosystem of the water body is in its ideal state for a healthy lifestyle.
2. **DISASTER**: This indicates that the ecology has deteriorated and is in poor condition. The fish can only survive for a short time. To prevent a tragedy, quick action is required.
3. **Just To Dwell (JTD):** This phrase must be understood to mean that the fish can live for a few weeks. Although it may not be in the best shape, the ecology must change for it to reach a state of ALIVE
4. **Just To Maintain (JTM)**: Under the JTD state, this circumstance exists. It follows that immediate action is required to prevent the situation from escalating into a DISASTER. For the environment to return to JTD or ALIVE and avert the catastrophe, be sure to do so. The decision can be thought over for a few days (three to five).

As shown in Table 3, four weights were established for the previously mentioned category.

Table 3. Result of Weights

Sl No	Result Outcome	Values
1	ALIVE	4
2	Just To Dwell (JTD)	3
3	Just To Maintain (JTM)	2
4	DISASTER	1

Based on these classification weights, a mathematical model was created to assess the fish's life conditions. We must first determine whether the parameters are directly related to any other parameter before we can examine the parameters. For the calculation of the Fish Classification Index (FCI), for example, temperature and dissolved oxygen (DO) are combined to generate a single weightage because they have an inverse relationship. Like Turbidity, there are some factors where other parameters do not directly affect it or where we can consider it to be a more independent variable. Its value is dependent on outside factors like the amount of rain falling or whether humans or animals have interfered with the water environment. Consequently, a different weighting has been given to this. This led to the derivation of the equation below.

Fish Classification Index (FCI) = DO & Temp + pH & UIA + Tur + Phosphate + Total BOD + Nitrate + Nitrite + Alkalinity
(1)

Whereas the range of values

DO & Temp = 1<Value>4

pH & UIA = 1<Value>4

Tur = 1<Value>4

Phosphate = 1<Value>4

Total BOD = 1<Value>4

Nitrate = 1<Value>4

Nitrite = 1<Value>4

Alkalinity = 1<Value>4

Therefore, the FCI = 1< Value > 32

According to our analysis and investigation, the FCI can have a value between 1 and 32. Below is a classification of our FCI value, which shows the likelihood of the fish surviving. Even yet, given the multiple outside factors that have an impact on the aquatic ecosystem, a prediction's error might still be as high as 20%. The computed values for each category follows because of this: -

Alive = FCI > 24 **(2)**

Just To Dwell = 24 <FCI > 16 **(3)**

Just To Maintain = 16 < FCI > 12
(4)

Disaster = 12 < FCI > 9 **(5)**

On the basis of the FCI value for each metric, we may evaluate the fish's life conditions. Additionally, the best values for each state were chosen, as shown in table 4.

Table 4. Value Classification Table

Sl No	Parameter	Disaster	Just To Maintain	Just To Dwell	Alive
1	Temperature in °C (T)	T > 38 or 10 <= T	10 < T<= 18	18< T<= 20	20<T<=35
2	Dissolved Oxygen in mg/l (DO)	DO > 12 DO < 2.5	2.5< DO <= 5	9< DO <=12	5< DO <=9
3	***pH***	3.5 <= pH >10	3.5< pH <=5 or 9 < pH <=10	5< pH <6.5	6.5 < pH <=9
4	Un-Ionized Ammonia in mg/l ***(UIA)***	UIA>0.5	0.1< UIA <0.3	0.05< UIA <0.09	0.05<= UIA
5	Turbidity (cm) ***(Tur)***	Tur < 12 or Tur >80	12<Tur<20 or 70<Tur<80	20<Tur <30 Or 60<Tur<70	30<Tur < 60
6	***Phosphate(mg/l)***	PO4>3	0.01<PO4 <0.02	0.02<PO4< 0.03	0.03 < Phos<3
7	***Nitrate***	0.01 < NO_3 NO_3 >100	0.05< NO_3 < 0.01 OR 50 < NO_3 < 100	0.01< NO_3 < 0.1 OR 4.5 < NO_3 < 49	0.1 < No_3 < 4.5
8	***BOD***	BOD>10	7<BOD<10	3<BOD<6	1<BOD<2
9	***Nitrite***	NO_2>0.2	0.1<NO_2<00.2	0.02<NO_2<0.1	0.02<NO_2
10	Alkalinity	20<Alk OR Alk > 300	20<Alk>35 OR 200<Alk<300	35<Alk<50 OR 100<Alk<200	25<Alk<100

4. Implementation

Using a smaller number of "summary indices" that are simpler to visualize and analyze, principal component analysis, or PCA, is a statistical technique that lets you condense the information contained in huge data tables. PCA serves as the foundation for multivariate data analysis using projection techniques. In order to identify trends, jumps, clusters, and outliers, it is crucial to describe a multivariate data table as a smaller number of variables.[25] With its great flexibility, PCA can analyze datasets with various characteristics, such as plurality, values that are lacking, categorized information, and inaccurate measurements. The objective is to separate the crucial information from the data and express it as a collection of summary indices known as primary components. As much data as feasible is gathered by each main component (PC) in a single index. The method generates a set of principal components that are not correlated, avoiding the inclusion of redundant data. The analysis generates a number of components, the first of which accounts for the majority of variance, the second of which accounts for the second largest amount of variance, and so forth.

Here an effort has been made to determine which of the parameters affects the Fish Classification Index's decision-making the most and by what proportion. A handful of data were considered on an Excel sheet. And using the Fish Classification Index, the suitable result was calculated. The results of the PCA analysis were obtained based on this classification as stated earlier, and the charts were plotted using Python on Jupyter Notebook Tool for Machine Learning, the seaborn library, and the numpy library. The excel csv file was considered as input to the program. The respective Eigen values and eigen vectors were generated as shown below.

Eigen Vectors:

-6.41839851e-03 6.93823779e-03 1.14723542e-02 1.43101574e-02

-5.77124747e-03 -8.80807137e-05]

[-7.06966490e-01 6.85194365e-01 -1.71777796e-01 -1.90536836e-02

-3.10697234e-03 3.12931493e-03 -1.84207112e-02 -2.08007393e-02

6.49919024e-03 1.65281146e-03]

[3.24939742e-04 -1.94682596e-02 -4.47987314e-02 -2.66964858e-02

-3.97938788e-03 -1.09994837e-02 5.81472735e-01 -6.95820321e-01

3.18677982e-01 -2.70059830e-01]

[7.30120999e-03 -9.86147896e-02 -2.58581524e-01 -1.23697692e-01

2.81732937e-02 1.65559733e-02 -7.38509527e-01 -6.01073128e-01

-1.73955749e-02 -7.90248879e-04]

[-5.33986913e-03 -1.56881877e-01 -6.55802093e-01 5.01318632e-03

6.46876945e-01 -2.85493894e-01 1.48015301e-01 1.48096973e-01

-2.65274899e-02 2.75054290e-02]

[-4.09695736e-03 4.52570261e-03 9.55376037e-02 -6.89025674e-01

-2.12045286e-01 -6.83068871e-01 9.53339993e-03 5.86414682e-02

2.82023538e-02 -1.44066841e-02]

[7.29164584e-03 -1.53517019e-01 -6.45446534e-01 -1.28650871e-01

-6.52058546e-01 2.45996541e-01 1.65135794e-01 1.05670180e-01

-1.05305505e-01 8.94827896e-02]

[1.30212009e-03 5.34987517e-03 8.38330207e-02 -6.99601530e-01

3.30750965e-01 6.15338914e-01 8.72621977e-02 3.38131404e-02

-5.55558511e-02 6.01394697e-02]

[-8.37524379e-04 2.91426812e-02 1.04448006e-01 4.78463164e-02

3.38800174e-02 -1.10654635e-01 2.43097070e-01 -3.39454391e-01

-7.23794085e-01 5.23789744e-01]

[1.24422405e-03 -4.99631959e-03 6.76975769e-03 1.39806830e-02

-1.48050058e-03 -7.51050709e-03 6.48680461e-03 -3.15814675e-02

5.98759994e-01 8.00075373e-01]]

Eigenvalues :

[1.30018143 3.20077212 2.200126678 1.30224926 0.23833426 0.36839322 0.66055608 0.42247246 0.1008730867 0.21132841]

The respective water quality parameter for the above eigenvalues should be interpreted in the order as listed. Temp, DO, pH, Amount of UIA, Turbidity, Phosphate, Nitrate, BOD, Alkalinity, Nitrite

5. Result And Graph Analysis

From Figure 2, the graph is got by plotting all the 10 components in correlation with Cumulative variance. As observed from the graph of Figure 3, 80% of the result is influenced by only the top four parameters, and the remaining is influenced by the other six parameters considered approximately. When considered individually the amount of influence of these four parameters is as follows in table 5.

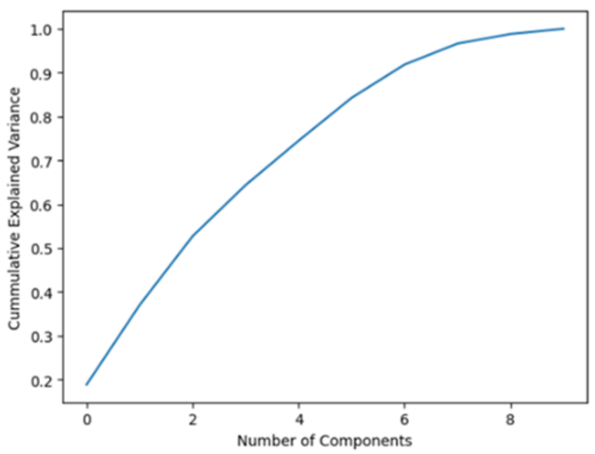

Fig.2. PCA Analysis Graph

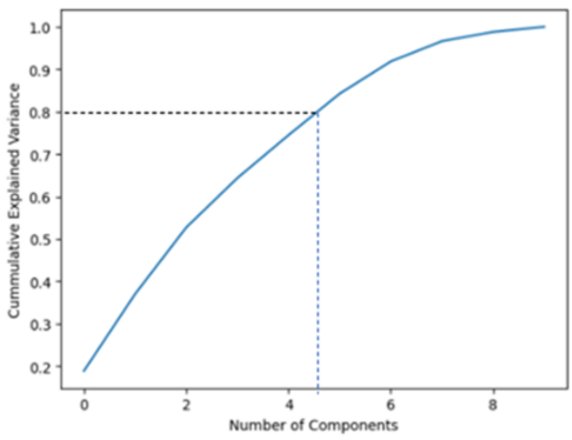

Fig.3. Indication of Significant Parameters

Table 5. Four Significant Parameters

Sl No	Parameter Name	Percentage of Impact
1	DO	32.00%
2	pH	22.01%
3	UIA	13.022%
4	Temperature	13.01%

Figure 4 below displays the overall distribution of the influence percentages of the major parameters on Fish Classification Index.

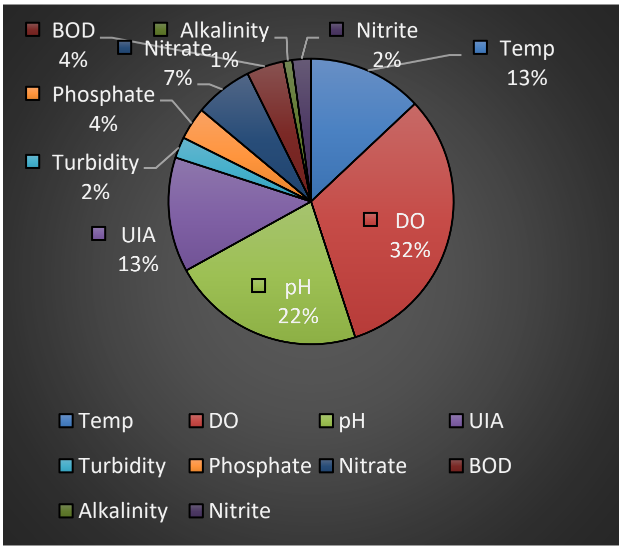

Fig.4. Components Impact & Distribution

The results of the investigation provide excellent information to engineers working on Internet of Things designs, aquaculture businesses, the fisheries division, and organizations in charge of lake development. There are fourteen to fifteen water quality metrics, thus IoT designers will be considering which parameter sensor to employ, which particular parameter is crucial, and how much of an impact a particular parameter might have on the health of the fish. For IoT designers, this conclusion would provide insight into which parameters have a significant impact and should be given top importance and also reduce the cost of their design. The values mentioned above can also provide excellent quantitative metrics, which help the relevant agencies think through the best course of action when a problem arises.

6. Conclusion

The study's findings offer a thorough analysis of how each factor affecting water quality affects the conditions under which fish can survive. These are the three outcomes attained: (1) The most important parameter that has to be checked in the water for the fish's healthy living conditions is given to the aquafarmer or lake agencies by the impact study. Ten different water quality factors were taken into account for the current study, but in that only four of them —DO, pH, UIA, and temperature — have a substantial effect on fish survival rates of about 80%. The remaining six water quality parameters have an impact on about 20% of the aquatic fish's living conditions. (2) The DO affects the most by 32%, temperature around 22%, UIA, and Temperature accounting for 13% approximately to the overall wellness of the fish. (3) This study offers crucial information for IoT designers to consider the appropriate sensors and support them with an efficient design to minimize potential costs. Additionally, it helps those organizations in charge of lake development to give the most important parameters high priority.

7. References and Footnotes

Author contributions

Dinesh M: Conceptualization, Methodology, Software, Field study **Prakash Sheelvanthmath:** Data curation, Writing-Original draft preparation, Validation., Field study **Anasuya N Jadagerimath:** Visualization, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare that this is original research and do not have any conflict.

References

[1] S. Ndlovu, S. Moyo, S. Nleya and S. Dube, "Precision Fish Farming Systems: A Mapping Study," 2022 1st Zimbabwe Conference of Information and Communication Technologies (ZCICT), Harare, Zimbabwe, 2022, pp. 1-5, doi: 10.1109/ZCICT55726.2022.10045996.

[2] H. F. Hawari and M. A. Hazwan, "Development of Iot

Monitoring System For Aquaculture Application," 2022 International Conference on Green Energy, Computing and Sustainable Technology (GECOST), Miri Sarawak, Malaysia, 2022, pp. 330-334, doi: 10.1109/GECOST55694.2022.10010661.

[3] H. Yuan et al., "Evaluation of Smart Environmental Protection Systems and Novel UV-Oriented Solution for Integration, Resilience, Inclusiveness and Sustainability," 2020 5th International Conference on Universal Village (UV), Boston, MA, USA, 2020, pp. 1-77, doi: 10.1109/UV50937.2020.9426222.

[4] S. Verma and Prachi, "Wireless Sensor Network application for water quality monitoring in India," 2012 NATIONAL CONFERENCE ON COMPUTING AND COMMUNICATION SYSTEMS, Durgapur, India, 2012, pp. 1-5, doi: 10.1109/NCCCS.2012.6412990.

[5] L. Yunfeng and Z. Tianpei, "A Design of Dissolved Oxygen Monitoring System Based on Nb-Iot," 2019 International Conference on Smart Grid and Electrical Automation (ICSGEA), Xiangtan, China, 2019, pp. 98-101, doi: 10.1109/ICSGEA.2019.00030.

[6] N. Shaghaghi, T. Nguyen, J. Patel, A. Soriano and J. Mayer, "DOxy: Dissolved Oxygen Monitoring," 2020 IEEE Global Humanitarian Technology Conference (GHTC), Seattle, WA, USA, 2020, pp. 1-4, doi: 10.1109/GHTC46280.2020.9342916.

[7] J. A. Cruz-Anchiraico, L. N. Mantari-Ramos, J. D. A. Cangalaya and D. Huamanchahua, "Design of an Automated System of PH and Water Level for an Aquaponic Module," 2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, NY, USA, 2022, pp. 0603-0608, doi: 10.1109/UEMCON54665.2022.9965721.

[8] Gholizadeh, Mohammad, et al. "A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques." Sensors, vol. 16, no. 8, MDPI AG, Aug. 2016, p. 1298. Crossref, https://doi.org/10.3390/s16081298.

[9] S. R. Khaire and R. M. Wahul, "Water Quality Data Analysis and Monitoring System in IoT Environment," 2018 3rd International Conference on Contemporary Computing and Informatics (IC3I), Gurgaon, India, 2018, pp. 326-330, doi: 10.1109/IC3I44769.2018.9007289.

[10] Kumar, Ajitesh. "Supervised and Unsupervised Learning Difference - Data Analytics." Data Analytics, 2 May 2023, vitalflux.com/dummies-notes-supervised-vs-unsupervised-learning.

[11] A. Singh, N. Thakur and A. Sharma, "A review of supervised machine learning algorithms," 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2016, pp. 1310-1315.

[12] F. Setievi, J. Natalia, T. R. Tjhang, I. S. Edbert and D. Suhartono, "A Comparative Study of Supervised Machine Learning Algorithms for Fake Review Detection," 2022 5th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 2022, pp. 306-312, doi: 10.1109/ISRITI56927.2022.10052860.

[13] N. Amruthnath and T. Gupta, "A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance," 2018 5th International Conference on Industrial Engineering and Applications (ICIEA), Singapore, 2018, pp. 355-361, doi: 10.1109/IEA.2018.8387124.

[14] R. B. Shrestha, M. Razavi and P. W. C. Prasad, "An Unsupervised Machine Learning Technique for Recommendation Systems," 2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA), Sydney, Australia, 2020, pp. 1-9, doi: 10.1109/CITISIA50690.2020.9371817.

[15] C. Yumeng and F. Yinglan, "Research on PCA Data Dimension Reduction Algorithm Based on Entropy Weight Method," 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), Taiyuan, China, 2020, pp. 392-396, doi: 10.1109/MLBDBI51377.2020.00084.

[16] Sarker, Iqbal H. "Machine Learning: Algorithms, Real-World Applications and Research Directions." SN Computer Science, vol. 2, no. 3, Springer Science and Business Media LLC, Mar. 2021. Crossref, https://doi.org/10.1007/s42979-021-00592-x.

[17] Y. Khan and C. S. See, "Predicting and analyzing water quality using Machine Learning: A comprehensive model," 2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT), Farmingdale, NY, USA, 2016, pp. 1-6, doi: 10.1109/LISAT.2016.7494106.

[18] E. Batur and D. Maktav, "Assessment of Surface Water Quality by Using Satellite Images Fusion Based on PCA Method in the Lake Gala, Turkey," in IEEE Transactions on Geoscience and Remote Sensing, vol. 57, no. 5, pp. 2983-2989, May 2019, doi: 10.1109/TGRS.2018.2879024.

[19] Xuedi Zhang, Jianhua Wu and Baode Song, "Application of principal component analysis in groundwater quality assessment," 2011 International Symposium on Water Resource and Environmental

Protection, Xi'an, 2011, pp. 2080-2083, doi: 10.1109/ISWREP.2011.5893671.

[20] M. M. Than, K. M. Yee, K. Lint, M. Han and T. W. Hnin, "Determining Spatial and Temporal Changes of Water Quality in Hlaing River using Principal Component Analysis," 2020 IEEE Conference on Computer Applications(ICCA), Yangon, Myanmar, 2020, pp. 1-7, doi: 10.1109/ICCA49400.2020.9022841

[21] M. Kumar, T. Singh, M. K. Maurya, A. Shivhare, A. Raut and P. K. Singh, "Quality Assessment and Monitoring of River Water Using IoT Infrastructure," in IEEE Internet of Things Journal, doi: 10.1109/JIOT.2023.3238123.

[22] R. K. Yadav, A. Jha and A. Choudhary, "IoT based prediction of water quality index for farm irrigation," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), Coimbatore, India, 2021, pp. 1443-1448, doi: 10.1109/ICAIS50930.2021.9395921.

[23] G. Hassan, M. E. Shaheen and S. A. Taie, "Prediction Framework for Water Quality Parameters Monitoring via Remote Sensing," 2020 First International Conference of Smart Systems and Emerging Technologies (SMARTTECH), Riyadh, Saudi Arabia, 2020, pp. 59-64, doi: 10.1109/SMART-TECH49988.2020.00029.

[24] T. Cheng, F. Harrou, Y. Sun and T. Leiknes, "Monitoring Influent Measurements at Water Resource Recovery Facility Using Data-Driven Soft Sensor Approach," in IEEE Sensors Journal, vol. 19, no. 1, pp. 342-352, 1 Jan.1, 2019, doi: 10.1109/JSEN.2018.2875954.

[25] “What Is Principal Component Analysis (PCA) and How It Is Used?” Sartorius, 18 Aug. 2020, [Online]. Available: www.sartorius.com/en/knowledge/science-snippets/what-is-principal-component-analysis-pca-and-how-it-is-used-507186.

[26] Singh, S. ., Bharti, A. K. ., Pandey, H. ., Yadav, R. K. ., Sharma, D. ., & Shanker, N. (2023). Towards Automated and Optimized Security Orchestration in Cloud SLA. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 215–222. https://doi.org/10.17762/ijritcc.v11i3.6339

[27] Ms. Madhuri Zambre. (2012). Performance Analysis of Positive Lift LUO Converter . International Journal of New Practices in Management and Engineering, 1(01), 09 - 14. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/3