



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Special Issue 2, April 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Machine Learning-Based Real-Time Water Quality Monitoring System

Nanda Jayalakshmi P^[1], Revathy V^[2]

PG Student, Department of Computer Science and Engineering, Grace College of Engineering, Thoothukudi.,
Tamil Nadu, India^[1]

Assistant Professor, Department of Computer Science and Engineering, Grace College of Engineering,
Thoothukudi., Tamil Nadu, India^[2]

ABSTRACT : Water quality parameter is of much importance in our day to day lives. Prediction of water quality will help to reduce water pollution and guard our human health. An intelligent process of monitoring the quality of water automatically detects the condition of water through IoT by processing sensors data and instantly provides notification to water analyst when the quality of water is abnormal. With the initiation of Machine to Machine Communication analyzing and communicating the data becomes simple and efficient. This work has advanced an “Machine Learning-based real-time water quality monitoring system” pertaining to lakes is being used in rural areas. The structure uses pH, Turbidity and Conductivity sensors for determining the water quality parameters about hydrogen ion and total dissolved solvents in the water. Likewise, K Means and Artificial Neural Network calculation has been utilized for anticipating the nature of water, with the assistance of prepared informational collection from various water tests.

KEYWORDS: Preprocessing, Segmentation, Extraction, classification

I. INTRODUCTION

The gift of water is one of the greatest natural resources humanity has ever received. However, contamination of water sources increased rapidly as a result of the rapid development of industrial companies and the abundance of human activities. In order to fix the underlying problems, it is crucial to monitor water quality in real time to detect any changes in quality parameters and guarantee that the water is fit for human consumption.

II. RELATED WORK

The concept of water quality monitoring system was first introduced by Fei yuan , yifan huang , xin chen , and en cheng et al [1], the method was introduced by A Biological Sensor System Using Computer Vision for Water Quality Monitoring. The type of sensors and number of sensor nodes provide high scalability ,low cost and low power consumption. Example based method proposed by kunyi zhang¹, ziqian dong [2], this method produced by microwave sensing of water quality. A high levels of nitrates in drinking water can be poisonous to human being which can also bring serious sickness and also even death. In [3], the author decomposes the IoT Based Realtime River Water Quality Monitoring System. In [4], the author developed Water Quality Monitoring System using IoT and Machine Learning. It is a network of devices connected to a sensor, Node microcontroller and a network which makes the communication possible among the devices. In[5],Energy Efficient Solutions in Wireless Sensor Systems for Water Quality Monitoring. With the usage of array, it helps us to collect more information in comparison with the single sensor system. The work in this paper is divided in two stages. 1. Preprocessing 2.Segmentation. Preprocessing is done by find out the original data and incorrect data are obtained from the given data and is used to find out the data cleaning using PH value. Segmentation Used to find out the pH level and water chemical reactions.

III. METHODOLOGY

The system effectually assist to screen the water quality parameter and analyze the quality of water at long distance without loss of data to avoid its life-threatening effects on Human beings, and on Animals. With the help of ML, they can have a continuous water quality monitoring system on various water parameters. All the ML data are stored and analyzed in the datasets using kaggle. Therefore, the proposed solution is very affordable (less cost). Unfortunately, it is impossible for human to monitor water's chemical composition frequently, which is an impossibly demanding task, let alone in real time. Thus, taking advantage of a highly efficient system can be an excellent solution. In this paper, we created a system called Real-time Intelligent Monitoring System for Water Quality, which enables a city to be responsive to potential outbreak of contamination and to protect city residents

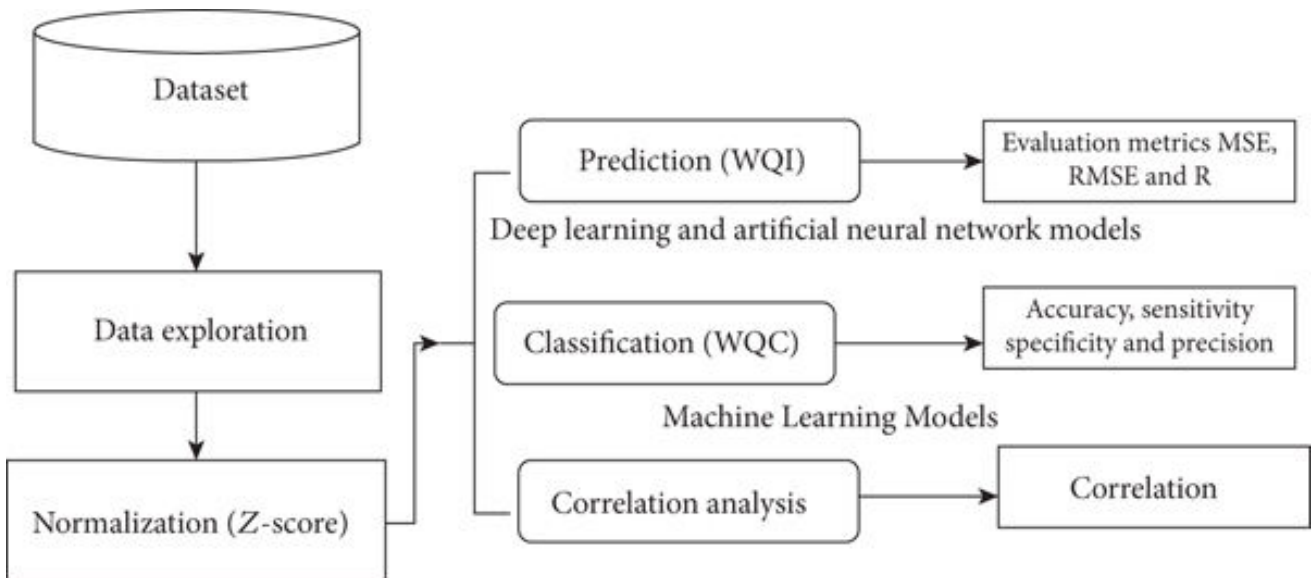


Fig 1 Architecture of water quality monitoring system

IV. EXPERIMENTAL RESULTS

Figures 2,3, 4 and 5 shows the results of water quality value from input dataset by using Adaptive filter algorithm. Figs. 2) shows the input dataset. (3) is finding the correct and incorrect data. (4)Preprocessing is used to find out the data cleaning using PH value. Through this, original data and incorrect data are obtained from the given data. Preprocessing requires finding the Mean value and Standard deviation value. It is found using Adaptive filter algorithm. (5)Segmentation is Used to find out the pH level and water chemical reactions.



```
In [11]: df.describe()
Out[11]:
```

	ph	Hardness	Solids	Chloramines	Sulfate	Conductivity	Organic_carbon	Trihalomethanes	Turbidity	Potability
count	3276.000000	3276.000000	3276.000000	3276.000000	3276.000000	3276.000000	3276.000000	3276.000000	3276.000000	3276.000000
mean	7.080795	196.369496	22014.092526	7.122277	333.775777	426.205111	14.284970	66.396293	3.966786	0.390110
std	1.469956	32.879761	8768.570828	1.583085	36.142612	80.824064	3.308162	15.769881	0.780382	0.487849
min	0.000000	47.432000	320.942611	0.352000	129.000000	181.483754	2.200000	0.738000	1.450000	0.000000
25%	6.277673	176.850538	15666.690297	6.127421	317.094638	365.734414	12.065801	56.647656	3.439711	0.000000
50%	7.080795	196.967627	20927.833607	7.130299	333.775777	421.884968	14.218338	66.396293	3.955028	0.000000
75%	7.870050	216.667456	27332.762127	8.114887	350.385756	481.792304	16.557652	76.666609	4.500320	1.000000
max	14.000000	323.124000	61227.196008	13.127000	481.030642	753.342620	28.300000	124.000000	6.739000	1.000000

```
In [12]: df.Potability.value_counts()
Out[12]:
0    1998
1    1278
Name: Potability, dtype: int64
```

Fig 2 – Input Dataset

```
In [1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

In [2]: df=pd.read_csv(r'water_potability.csv')
df.head()
Out[2]:
```

	ph	Hardness	Solids	Chloramines	Sulfate	Conductivity	Organic_carbon	Trihalomethanes	Turbidity	Potability
0	NaN	204.890455	20791.318981	7.300212	368.516441	564.306654	10.379783	86.990970	2.963135	0
1	3.716080	129.422921	18630.057958	6.635246	NaN	592.885359	15.180013	56.329076	4.500656	0
2	8.009124	224.236259	19908.541732	9.275884	NaN	418.806213	16.868837	66.420093	3.055934	0
3	8.316766	214.373394	22018.417441	8.059332	356.889136	363.266516	18.436524	100.341674	4.628771	0
4	9.092223	181.101509	17978.986339	6.548600	310.135738	398.410613	11.558279	31.997993	4.075075	0

```
In [3]: df.shape
Out[3]: (3276, 10)
```

Fig 3 – Finding the correct data

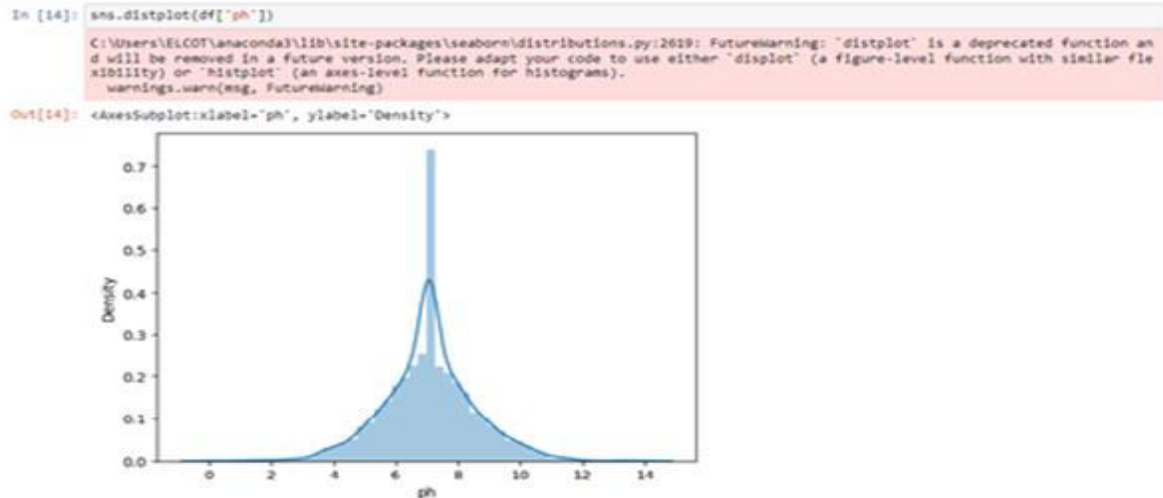


Fig 4 - Preprocessing

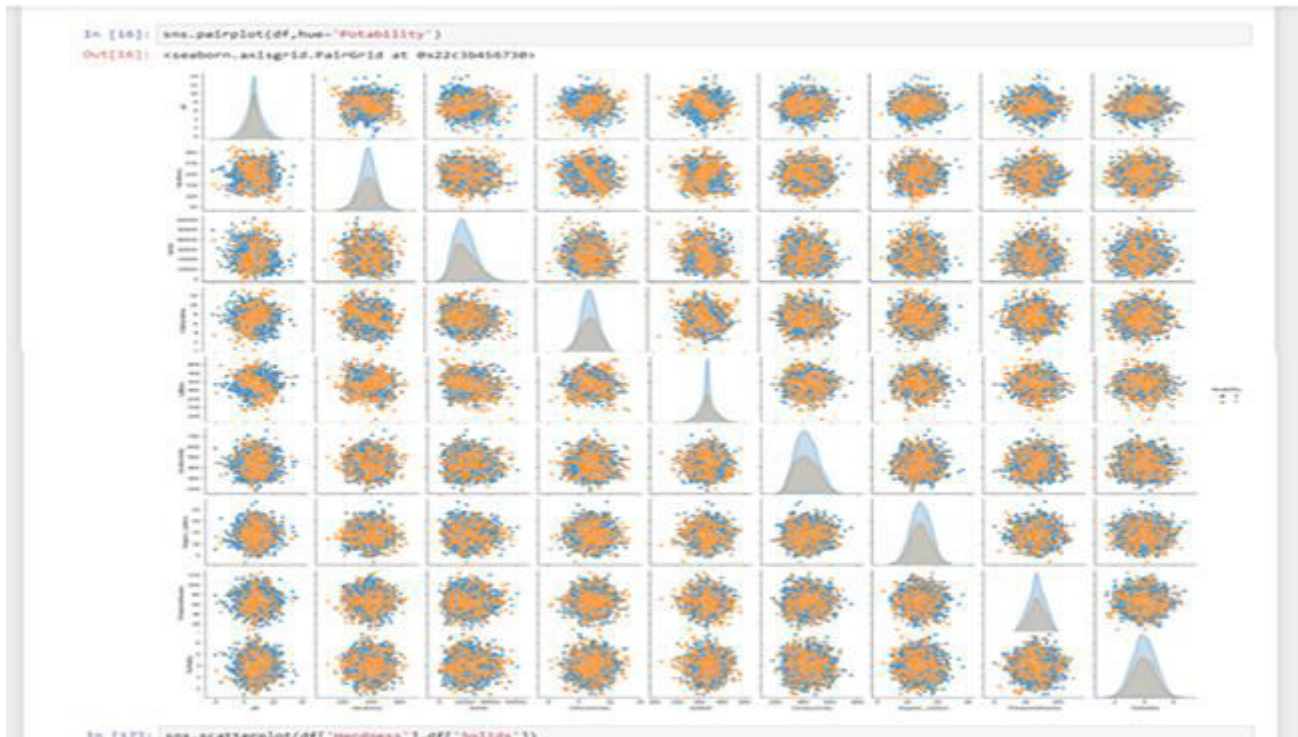


Fig 5- Segmentation

V. CONCLUSION

The initial dataset is preprocessed with a statistical stratified sampling strategy. Experimental results show that the proposed model with stratified sampling outperforms MLR, ANNs, ANFIS, EANFIS, and WT-ANFIS models. This demonstrates that the proposed model, with stratified sampling and a general-purpose input parameter selection method, is reliable to model the quality of parameters related to salinity. This model can be applied to reduce the number of parameters monitored to lower the cost associated with monitoring the quality of wastewater discharges.

REFERENCES

1. Y. Hirabayashi *et al.*, "Global flood risk under climate change," *Nature Climate Change*, vol. 3, no. 9, pp. 816–821, 2013.
2. MunichRE, "Risks from floods, storm surges and flash floods," 2019. [Online]. Available: <https://www.munichre.com>
3. P. Brivio, R. Colombo, M. Maggi, and R. Tomasoni, "Integration of remote sensing data and GIS for accurate mapping of flooded areas," *Int. J. Remote Sens.*, vol. 23, no. 3, pp. 429–441, 2002.
4. A. Goffi, D. Stroppiana, P. A. Brivio, G. Bordogna, and M. Boschetti, "Towards an automated approach to map flooded areas from Sentinel-2 MSI data and soft integration of water spectral features," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 84, 2020, Art. no. 101951.
5. Y. Du, Y. Zhang, F. Ling, Q. Wang, W. Li, and X. Li, "Water bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band," *Remote Sens.*, vol. 8, no. 4, p. 354, 2016, Art. no. 354.
6. R. Brakenridge and E. Anderson, "Modis-based flood detection, mapping and measurement: The potential for operational hydrological applications," in *Transboundary Floods: Reducing Risks Through Flood Management*. Berlin, Germany: Springer, 2006, pp. 1–12.



7. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. Int. Conf. Mach. Learn.*, vol. 30, no. 1, 2013
8. D. Amitrano, G. Di Martino, A. Iodice, D. Riccio, and G. Ruello, "Unsupervised rapid flood mapping using Sentinel-1 GRD SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3290–3299, Jun. 2018.
9. F. Gascon *et al.*, "Copernicus Sentinel-2A calibration and products validation status," *Remote Sens.*, vol. 9, no. 6, 2017, Art. no. 584.
10. L. Pulvirenti, N. Pierdicca, M. Chini, and L. Guerriero, "An algorithm for operational flood mapping from synthetic aperture radar (SAR) data using fuzzy logic," *Natural Hazards Earth Syst. Sci.*, vol. 11, no. 2, pp. 529–540, 2011.
11. M. C. Hansen and T. R. Loveland, "A review of large area monitoring of land cover change using landsat data," *Remote Sens. Environ.*, vol. 122, pp. 66–74, 2012.
12. S. Schlaffer, P. Matgen, M. Hollaus, and W. Wagner, "Flood detection from multi-temporal SAR data using harmonic analysis and change detection," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 38, pp. 15–24, 2015.
13. S. Long, T. E. Fatoyinbo, and F. Policelli, "Flood extent mapping for Namibia using change detection and thresholding with SAR," *Environ. Res. Lett.*, vol. 9, no. 3, 2014, Art. no. 035002.
14. M. Rahnemoonfar, T. Chowdhury, A. Sarkar, D. Varshney, M. Yari, and R. Murphy, "Floodnet: A high resolution aerial imagery dataset for post flood scene understanding," *IEEE Access*, vol. 9, pp. 89644–89654, 2020.
15. M. Berger, J. Moreno, J. A. Johannessen, P. F. Levelt, and R. F. Hanssen, "ESA's Sentinel missions in support of Earth system science," *Remote Sens. Environ.*, vol. 120, pp. 84–90, 2012



SJIF: 8.423



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INTERNATIONAL CENTRE



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details