



Article Title: Machine Learning Based Realtime Water Quality Monitoring System

Machine Learning Based Realtime Water Quality Monitoring System

Nanda Jayalakshmi P¹, Revathy V²

¹ME scholar, Department of Computer Science and Engineering, Grace College of Engineering, Thoothukudi-628 004, India.

²Assistant Professor, Department of Computer Science and Engineering, Grace College of Engineering, Thoothukudi-628 004, India.

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. This work has advanced an “Machine Learning based real-time water quality monitoring system” pertaining to lakes is being used in rural areas. All the organisms that use it are affected by the waste generated in this water. Water quality monitoring system is to identify the level of water and finding ways to correct the problems in it. Water quality refers to the chemical, physical and biological characteristics of water. It is a measure of the condition of water relative to the requirements of one or more biotic species and or to any human need or purpose. It is most frequently used by reference to a set of standards against which compliance can be assessed. The most common standards used to assess water quality relate to health of ecosystems, safety of human contact and drinking water. An intelligent process of monitoring the quality of water automatically detects the condition of water through Machine Learning by processing sensors data and instantly provides notification to water analyst when the quality of water is abnormal. The structure uses Determination of pH, Color, Temperature, Carbon monoxide, Conductivity, Fecal coliform. Likewise, ANN, Adaptive filter algorithm, Support Vector Machine, Naïve Bayes, Random forest calculation has been utilized for anticipating the nature of water, with the assistance of prepared informational collection from various water tests.

Keywords: Water quality, Drinking water, pH, Water quality monitoring

1 Introduction

The project was based on testing the quality of water. Water quality refers to the chemical, physical and biological characteristics of water. It is a measure of the condition of water relative to the requirements of one or more biotic species and or to any human need or purpose. It is most frequently used by reference to a set of standards against which compliance can be assessed. The most common standards used to assess water quality relate to health of ecosystems, safety of human contact and drinking water. 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



Article Title: Machine Learning Based Realtime Water Quality Monitoring System

water quality in real time to detect any changes in quality parameters and guarantee that the water is fit for human consumption.

2 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. Exemplar 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.

3 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



Article Title: Machine Learning Based Realtime Water Quality Monitoring System

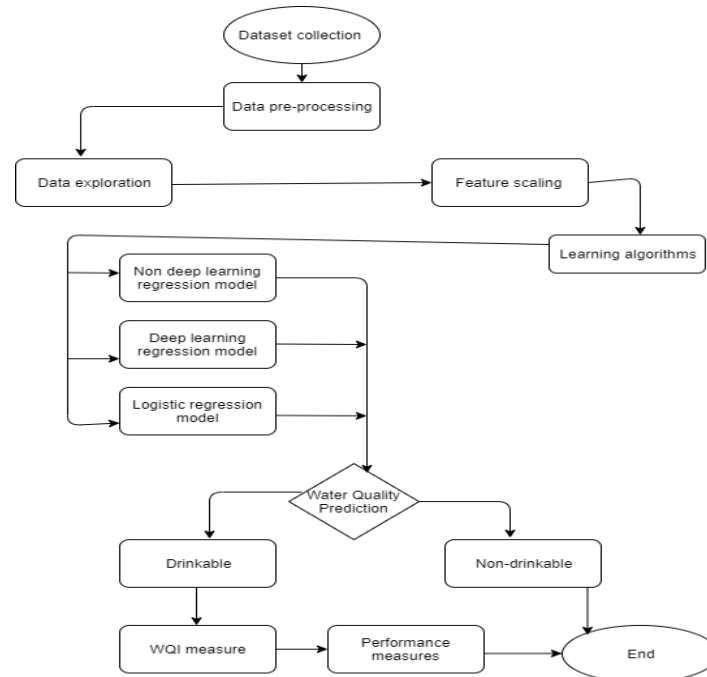


Figure 1: Architecture of water quality monitoring system

4 Experimental Results

Figures 2, 3, 4 and 5 shows the results of water quality value from input dataset by using Adaptive filter algorithm. Figures. 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)Problem solving method 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	196.389496	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	80.824054	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

```

The chart displays the number of publications per year for various countries from 2006 to 2014. The y-axis represents the number of publications, ranging from 0 to 2,500,000. The x-axis represents the year. Most countries show very low publication counts, while the USA shows a significant increase starting in 2011, peaking in 2012 and 2014.

Country	2006	2007	2008	2009	2010	2011	2012	2013	2014
usa	0	0	0	0	0	0	~2,100,000	~1,100,000	~2,600,000
china	0	0	0	0	0	0	~100,000	~100,000	~100,000
uk	0	0	0	0	0	0	~100,000	~100,000	~100,000
germany	0	0	0	0	0	0	~100,000	~100,000	~100,000
france	0	0	0	0	0	0	~100,000	~100,000	~100,000
india	0	0	0	0	0	0	~100,000	~100,000	~100,000
australia	0	0	0	0	0	0	~100,000	~100,000	~100,000
canada	0	0	0	0	0	0	~100,000	~100,000	~100,000
italy	0	0	0	0	0	0	~100,000	~100,000	~100,000
spain	0	0	0	0	0	0	~100,000	~100,000	~100,000
other countries	0	0	0	0	0	0	~100,000	~100,000	~100,000

Quality Parameter	Number of Parameters
Fluoride	10
Salinity	3

5 Conclusion

Page 372 of 374



Article Title: Machine Learning Based Realtime Water Quality Monitoring System

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. Fei Yuwan; Yifan Huang; Xin Chen; En cheng, Year: 2018, "A biological sensor system using computer vision for water quality monitoring", Fundamental research funds and central univ, November 2018.
2. L. Metcalf; H. P. Eddy, Year: 2014, "Wastewater Engineering: Treatment and Resource Recovery, 5th Edition. Mc Graw Hill, Chapter 3.
3. J. Wan; M. Huang; Y. Ma; W. Guo; Y. Wang; H. Zhang; W. Li; X. Sun, "Prediction of effluent quality of a paper mill wastewater treatment using an adaptive network-based fuzzy".
4. H. Zare, Year: 2014, "Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters", Journal of Environmental Health Science & Engineering, Vol: 12, no: 1, pp. 1 – 8.
5. L. Li; P. Jiang; H. Guang; L. Dong; G. Wu; H. Wu, Year: 2019, "Water quality prediction based on recurrent neural network and improved evidence theory: a case study of Qiantang River, China", Environmental Science and Pollution Research, Vol: 26, no: 19, pp. 19879 – 19896.
6. J. –S. R. Jang, Year: 1993, "ANFIS: Adaptive- Network-Based Fuzzy Inference Systems", IEEE Transactions on Systems, Man, and Cybernetics, Vol: 23, no: 3, pp. 665 – 685.
7. A. El-Shafie Najah; O. A. Karim; A. H. El-Shafie, Year: 2013, "Performance of ANFIS versus MLP- NN dissolved oxygen prediction models in water quality monitoring", Environmental Science and Pollution Research, Vol: 21, no: 3, pp. 16581670.
8. L. Zhang; G.X. Zhang; R.R. Li, Year: 2015, "Water quality analysis and prediction using hybrid time series and neural network models", Journal of Agricultural Science and Technology, Vol: 18, no: 4, pp. 975 – 983.
9. M. J. Alizadeh; M. R. Kavianpour, Year: 2015, "Development of wavelet-ANN models to predict water quality parameters in Hilo Bay, Pacific Ocean", Marine Pollution Bulletin, Vol: 98, no: 1, pp.171 – 178, 2015.
10. A.G. Mohammad; K. Reza; C. Kwok-Wing; S. Shahabuddin; T. G. Pezhman, Year: 2011, "forecasting pan evaporation with an inteligital neural network quantum -behaved inference system", Applied Soft Computing, Vol: 11, no: 3, pp.3238 – 3246.
11. W. Deng; G. Wang; X. Zhang, Year: 2015, "A novel hybrid water quality time series prediction



Article Title: Machine Learning Based Realtime Water Quality Monitoring System

method based on cloud model and fuzzy forecasting”, Chemometrics and Intelligent Laboratory Systems, Vol: 149, pp. 39 – 49.

12. P. Singh, Year: 2018, “Rainfall and financial forecasting using fuzzy time series and neural network based model”, International Journal of Machine Learning and Cybernetics, Vol: 9, no: 3, pp. 491 – 506.

13. A. K. Kadam; V. M. Wagh; A. A. Muley; B. N. Umrikar; R. N. Sankhua, Year: 2019, “Prediction of water quality index using artificial neural network and multiple linear regression modelling approach in Shivganga River basin, India”, Modeling Earth Systems and Environment, Vol: 5, no: 3, pp. 951 – 962.

14. A. Sarkar; P. Pandey, Year: 2015, “River Water Quality Modelling Using Artificial Neural Network Technique”, Aquatic Procedia, Vol: 4, pp. 1070 – 1077.

15. A. P. Piotrowski; M. J. Napiorkowski; J. J. Napiorkowski; M. Osuch, Year: 2015, “Comparing various artificial neural network types for water temperature prediction in rivers”, Journal of Hydrology, Vol: 529, no: 1, pp. 302 – 315.

16. A. Najah; A. El-Shafie; O. A. Karim; A. H. El-Shafie, Year: 2012, “Application of artificial neural networks for water quality prediction”, Neural Computing and Applications, Vol: 22, no: 1, pp. 187 – 201.