Water Environment Quality Assessment and Information System Based on Neural Network and Logistic Regression

Cherif Guerroudj[*]

Tokyo Polytechnic University, 1583 Iiyama, Atsugi 243-0297, Kanagawa, Japan

[*]*corresponding author*

Abstract: Water resources have always been an indispensable and important resource in the development of modern society. At the same time, water resources play an important role in the high-quality development of social economy and the daily life and work of residents. However, with the gradual deepening of the brutal development model of heavy industry and manufacturing, the degree of damage to water resources in different regions is also deepening. The further development of this water pollution problem has had certain restrictions and impacts on the development of social economy and the daily life and work of residents in different regions. Therefore, the pollution analysis and treatment mode of water environment has become increasingly important. After analyzing the pollution of the water environment in different regions, it can not only help the relevant staff to put forward a more reasonable pollution control model of the water environment, but also help the social economy to achieve better development. In the assessment of the pollution of different water environments, it is necessary to first analyze the quality of the water environment in the basin and comprehensively evaluate the water environment in combination with the consideration of various environmental factors of the regional environment. In this paper, a new water environment quality evaluation and information system is constructed by using neural networks (NN) and logical regression algorithm model. This new water environment quality evaluation and information system mainly uses NN and logical regression algorithm, and its data analysis and processing ability can perform efficient operations on various data in different water environments, thus greatly improving the efficiency of water environment quality evaluation. On the other hand, NN and logistic regression algorithm model can also play a positive role in improving the accuracy of water environment assessment. Finally, the performance difference between the water environment quality assessment and information system combined with NN and logistic regression algorithm and the existing water environment assessment system is analyzed, and it is determined that the performance of this new system in many aspects has been improved by about 18% on average.

1. Introduction

The quality assessment of water environment is not only a common method for relevant staff to manage the environment of a certain region, but also an important basis for understanding the development law of the environmental quality of the region, which can play a positive role in formulating the water pollution control model applicable to the region. Therefore, this paper mainly studies the quality evaluation and information system of water environment, and proposes a new system through the integration of NN and logistic regression algorithm.

Some water environmental protection researchers first study the current pollution situation in the water environment, hoping to determine some reasons that have a greater impact on the quality of the current water environment through these studies. Tang Yankui explored the category and content of chemical and physical factors in the pollutants in the target waters, which played a good role in the pollution control and overall water quality assessment of the target waters [1]. Singh Nirala explored the effect of an electrocatalytic chemical method in the pollution control of the target waters, and also determined the content of chemical pollutants in the target waters through this method [2]. Mekonnen Mesfin M analyzed the impact of the content of a pollutant in the target freshwater water environment on the quality of the water area, and determined that the content of the pollutant in the target water environment has a direct impact on the water quality [3]. Singh Upma explored the impact of industrial production waste on water quality after it was directly discharged into the water, and determined that the pollutants in industrial production have a direct impact on water quality [4].

Liu Yi studied the relationship between the degree of water pollution, water quality and social and economic development. Through the study of the relationship between water pollution and water environmental quality, it was determined that there was a direct relationship between the three [5]. Yan Yan explored the relationship between the degree of surface pollution in the target water area and the quality of the water area. Through the study of the degree of surface pollution in the water area, it was determined that the main factors affecting the quality of the water environment were mainly limited by the degree of surface water pollution [6]. He Mingjing explored the relationship between the pollution degree of water environment and the sustainable development of society, and determined that the pollution of water environment, that is, the quality of water environment, has a direct impact on the sustainable development of society [7]. However, most of these studies on the pollution status in the water environment are limited to the analysis of pollutants, and cannot study the factors that affect the quality of the water environment.

Another part of researchers have studied the assessment of water environment, namely water quality analysis, hoping to get a more complete water quality assessment model. Celekli Abuzer explored the impact of the content of diatom and other substances in a water area on the water quality assessment of that water area. Through in-depth research, it was determined that the content of diatom substances has a direct impact on the water quality assessment [8]. Ernazarovna Makhmudova Dildora explored the role of water quality assessment model in water safety in the target waters, and determined the role of water quality assessment model in safe water use in the region [9]. Solangi Ghulam Shabir explored the performance of water quality index in water quality assessment and determined that water quality index can play a positive role in water quality assessment model [10]. Camara Moriken explored the relationship between the utilization of land resources and the quality of water environment, and determined that the higher the land utilization rate, the greater the impact on the quality of water environment [11]. Bisht Anil Kumar explored the existing water environment quality analysis model and determined the feasibility of optimizing some emerging information technologies in it [12]. However, these research results are very fragmented and cannot support the construction of a complete water quality analysis and analysis

model. Therefore, more in-depth research is needed.

First of all, this paper deeply analyzes the operation process and final effect of the existing water environment quality evaluation system, discusses the causes of the defects in the operation process of the existing water environment quality evaluation system, and determines how to improve the defects. In addition, the advantages of the new water environment quality assessment and information system combined with NN and logistic regression algorithm model are analyzed, and it is determined that the new water environment quality assessment and information system has better improved the water environment quality analysis.

2. NN and Logistic Regression

With the continuous development of various modern information technologies, a variety of emerging information technologies have been deeply applied in many fields of society, and have had a great positive impact on the operation mode of many fields. NN is a kind of emerging information technology, which can simulate human learning ability and intelligence to a certain extent through in-depth analysis of human brain structure [13]. Through this simulation of human intelligence, NN has been widely used in many fields such as industry, finance and computer. This paper is to explore the application of NN in the water environment quality evaluation and information system, and determine the reliability of the performance of this new water environment quality evaluation and information system. The operation flow of general NN is shown in Figure 1.
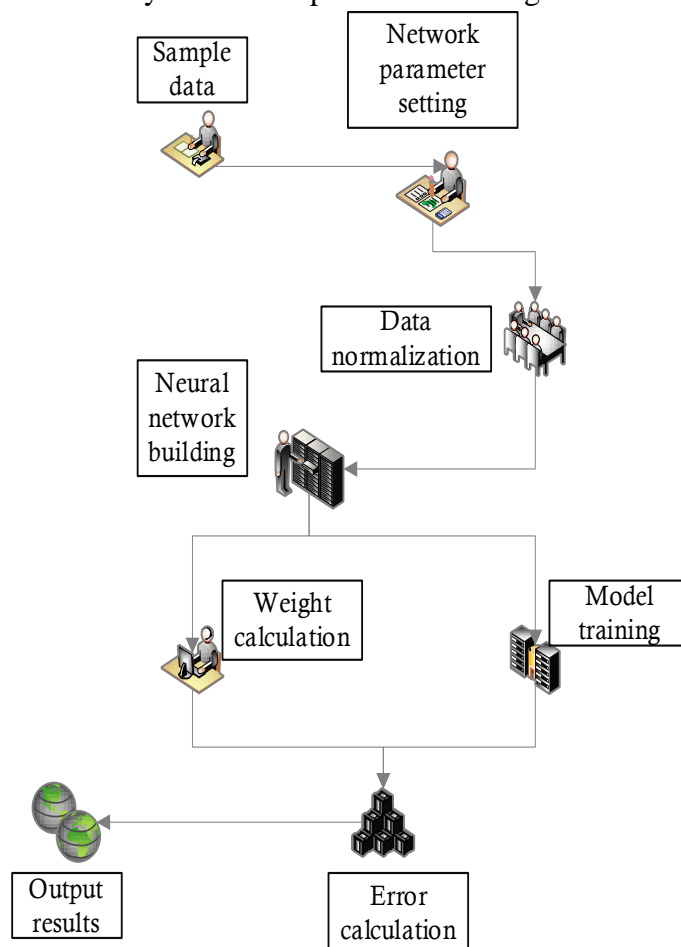

Figure 1. Schematic diagram of the operation flow of NN under normal circumstances

Like NN, the logistic regression model is a new technology in the field of artificial intelligence, which is popular in recent years, and both technologies have played a certain role in promoting the development of social economy. Logistic regression is mainly used to predict the development of data through the analysis of independent variables and dependent variables in the sample data set [14]. At the same time, the logical regression algorithm model can also explore a linear relationship between the sample data, and has good operational efficiency. Thanks to such characteristics, the logical regression algorithm model and NN can play a better role in the quality evaluation of water environment. On the one hand, it can greatly improve the efficiency of the data analysis process of water environment. On the other hand, it is to improve the accuracy of the calculation results in the water environment quality assessment, so as to make the water environment quality assessment perform better. The operation flow of the logical regression algorithm is shown in Figure 2.
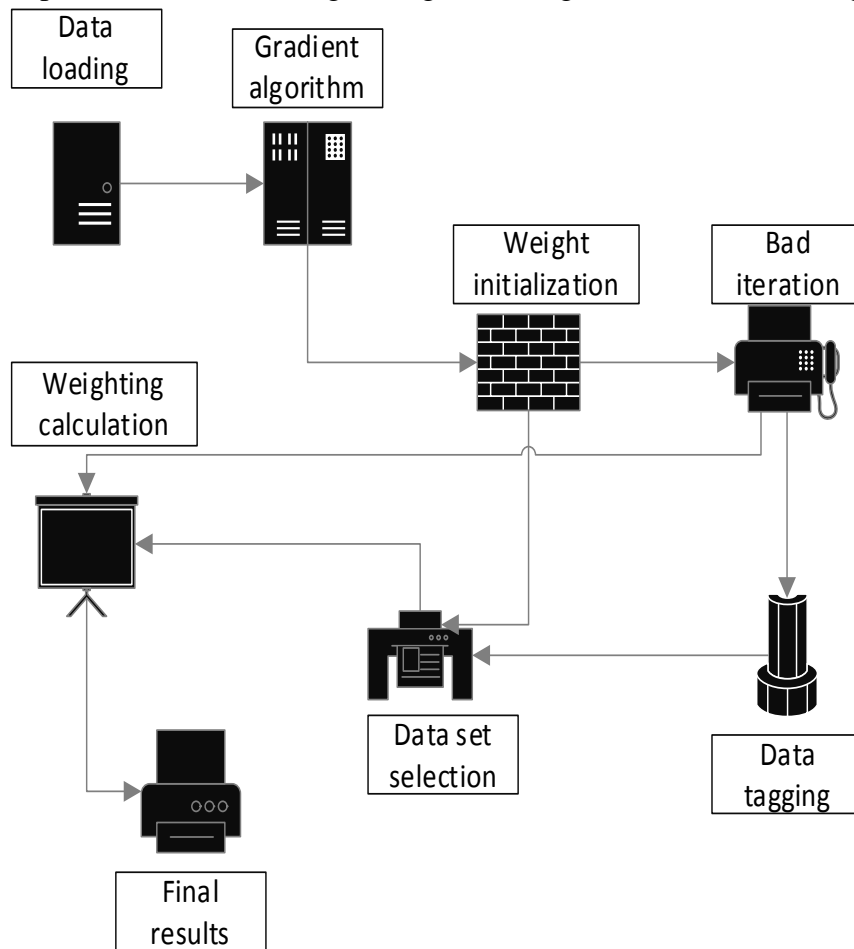

Figure 2. Schematic diagram of the operation flow of the logistic regression algorithm

3. Water Environment Quality Assessment and Information System

The quality assessment of water environment has always been one of the most important tasks in the development and use of water resources, and also an important research topic in the field of environmental protection [15]. In order to deeply apply NN and logistic regression algorithm model in water environment quality evaluation and information system, this paper first determines the topological structure of the evaluation model. The construction of this topological structure can not only have a direct impact on the accuracy of data operation in the water environment quality evaluation and information system, but also help NN and logical regression algorithm to be more

deeply integrated into the water environment quality evaluation and information system. By integrating NN and logistic regression algorithm with the water environment quality assessment model, this new model can not only generate learning samples for model training based on the water quality standards within the specified range, so as to automatically evaluate the water environment quality, but also have the adaptive and self-organizing ability of NN, so as to automatically evaluate the water environment quality of the target waters. The operation flow of the water environment quality assessment system combining NN and logistic regression algorithm is shown in Figure 3.
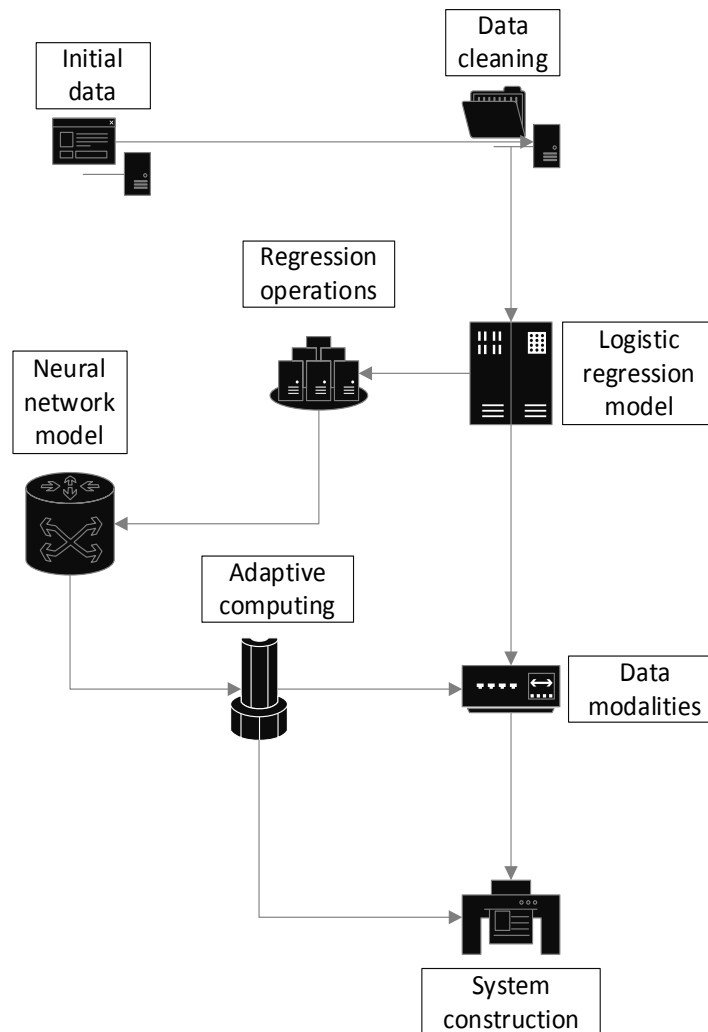

Figure 3. Schematic diagram of the operation process of the water environment quality evaluation system combining NN and logistic regression algorithm

4. Logic Regression Algorithm

At present, many mathematical algorithm models such as index method and fuzzy evaluation method are generally used to evaluate the quality of water environment, but the water environment quality analysis model formed by these algorithm models also has various shortcomings, which are further amplified with the continuous development of social economy. Therefore, researchers in the field of water resources protection have begun to discuss the optimization scheme for the deficiencies in the existing water environment quality assessment model. In reality, the more easily developed and relatively obvious improvement scheme is the integration of modern information

technology. This deep combination of modern information technology can significantly improve the efficiency and performance of the existing water environment quality assessment methods. This paper verifies the feasibility and reliability of a water environment quality assessment and information system combining NN and logical regression algorithm, and determines the superior performance of the algorithm in NN and logical regression algorithm model in the application of water environment quality assessment and information system.

First of all, a linear model is constructed by using the relevant formula to analyze the multiple types of data that affect the water environment quality, so as to determine the comprehensive quality of the water environment. The main construction formula of linear model is shown in (1).

$$f(x) = \frac{1}{1+e^{-a_w x}} \tag{1}$$

a_w mainly represents the weight of sample data. Then, the loss value in the process of data calculation is calculated, and the formula is shown in (2).

$$P = y_i(1-p)^{-y_i} \tag{2}$$

y_i represents the real output value and p represents the total probability. Then, the gradient descent is used to further calculate the sample data, and the formula is shown in (3).

$$y = \frac{e^{-a_w x}}{1+e^{-a_w x}} \tag{3}$$

These are some algorithm models used in this paper. Through these algorithm models, the existing water environment quality can be optimized, thus improving the overall work efficiency.

5. Experiment on the New Model of Water Environmental Quality Assessment and Information System

With the continuous progress of various modern information technologies, industry, manufacturing and other related industries have achieved full development, but the development of heavy industry in industry has also caused a greater negative impact on the sustainable development of water environment in different regions. In addition, people's uncontrolled use and waste of water resources in modern society have also led to serious problems such as insufficient supply and demand of water resources in some regions. Therefore, relevant departments pay more attention to the sustainable development of water environment. This emphasis is also reflected in the development of the quality assessment system for water environment in different regions. The quality assessment system for water environment in different regions is the basic workflow of the water environment prevention and control model. The quality assessment of water environment can help relevant staff to understand the water environment in depth. With the deepening of the complexity of water pollution in various regions, people's attention to the evaluation system of water environment quality is also gradually increasing. Various emerging information technologies are beginning to be integrated into the evaluation system of water environment quality, providing more development ideas for the quality analysis and evaluation model of water environment. In this paper, NN and logistic regression algorithms are used to optimize the existing water environment quality evaluation system.
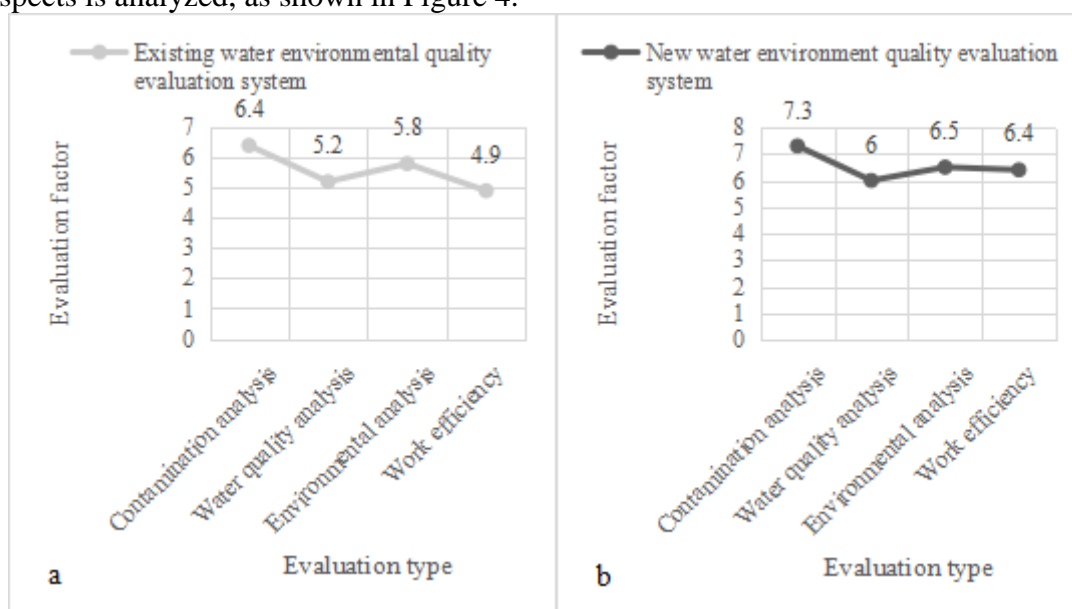
The first is to analyze the evaluation effects of various water environment quality assessment models in water resources, so as to analyze the existing water environment quality assessment models, as shown in Table 1.

Table 1. Performance of the quality evaluation model of the water environment

	Water quality index evaluation	Fuzzy integrated evaluation
Acidity and alkalinity	7.2	7.4
Ammonia nitrogen	5.9	6.4
Dissolved oxygen	6.7	7.3

Through the analysis of the evaluation performance of various substances in the two existing water environment quality evaluation models, it is determined that the fuzzy comprehensive evaluation performs relatively well in the evaluation of pH, ammonia nitrogen content and dissolved oxygen of water quality. On the one hand, the fuzzy comprehensive evaluation model can analyze various contents, and finally comprehensively evaluate the quality of water environment. Therefore, this evaluation model has higher accuracy.

Then, the performance of the new water environment quality assessment and information system combined with NN and logistic regression and the existing water environment quality assessment in many aspects is analyzed, as shown in Figure 4.

a. Performance diagram of existing water quality assessment system

b. New water environment quality evaluation system performance diagram

Figure 4. Schematic diagram of the performance of the new water environment quality evaluation and information system proposed in this paper and the existing water environment quality evaluation system

Finally, the performance of the existing water environment quality assessment model in Figure 4a in various assessment types is analyzed, and it is determined that the performance of the existing water environment assessment model in water quality analysis and work efficiency is relatively poor. On the other hand, it is to analyze the performance of the water environment quality assessment model and information system proposed in this paper in combination with NN and logistic regression algorithm model in Figure 4b when evaluating the water environment quality, and determine that the new water environment quality assessment model and information system have been improved in many aspects. Finally, the performance difference between the new water

environment quality assessment and information system and the existing water environment quality assessment model is analyzed comprehensively, and it is determined that the multiple performance of this new water environment quality assessment model has been improved by about 18%.

6. Conclusion

The sustainability of water resources and water environment has always been one of the focus issues in the field of environmental protection in the current society. The comprehensive evaluation of the quality of water environment can not only help the relevant staff understand the development rules of the water environment quality of the target water area, but also help the timely collection and analysis of the pollution situation of the target basin. At the same time, this workflow can also assist relevant researchers to quickly understand the source of pollution, the distribution of pollutants and the degree of pollution in water pollution, so as to quickly find out the reasons that affect the quality of the target water area and control the environment of the target water area. This work of comprehensive analysis of the quality of water environment can also summarize the water environment protection work in the target area and evaluate the effect of water environment protection and treatment over a period of time, which is also an important part of maintaining the high-quality development of social economy. However, the existing water environment quality evaluation system is still in a mode of manual collection and analysis, which also makes the existing water environment quality evaluation system have low efficiency and low accuracy of the final results. In this paper, NN and logistic regression algorithm are used to improve these deficiencies, such as automatic recording of water environmental quality evaluation data in history, comprehensive analysis and comparison of water environmental quality in each year, and so on. This new water environment quality assessment and information system, which combines NN and logistic regression algorithm model, makes the water environment quality assessment model more and more information-based.

Funding

Data Availability

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Conflict of Interest

The author states that this article has no conflict of interest.

References

[1] Yankui Tang, et al. Emerging pollutants in water environment: Occurrence, monitoring, fate, and risk assessment. Water Environment Research. (2019) 91(10): 984-991. https://doi.org/10.1002/wer.1163

[2] Singh Nirala, Bryan R. Goldsmith. Role of electrocatalysis in the remediation of water pollutants. ACS Catalysis (2020) 10(5): 3365-3371. https://doi.org/10.1021/acscatal.9b04167

[3] Mekonnen Mesfin M., Arjen Y. Hoekstra. Global anthropogenic phosphorus loads to freshwater and associated grey water footprints and water pollution levels: A high‑resolution global

study. *Water resources research.* *(2018)* *54(1):* *345-358.* *https://doi.org/10.1002/2017WR020448*

[4] *Singh Upma. Water Pollution due to Discharge of Industrial Effluents with special reference to Uttar Pradesh, India-A review. International Archive of Applied Sciences and Technology. (2018) 9(4): 111-121.*

[5] *Yi Liu, Liyuan Yang, Wei Jiang. Qualitative and quantitative analysis of the relationship between water pollution and economic growth: a case study in Nansi Lake catchment, China. Environmental Science and Pollution Research. (2020) 27(4): 4008-4020. https://doi.org/10.1007/s11356-019-07005-w*

[6] *Yan Yan. Ecological risk assessment from the viewpoint of surface water pollution in Xiamen City, China. International Journal of Sustainable Development & World Ecology. (2018) 25(5): 403-410. https://doi.org/10.1080/13504509.2017.1422567*

[7] *Mingjing He. Waste-derived biochar for water pollution control and sustainable development. Nature Reviews Earth & Environment. (2020) 3(7): 444-460.*

[8] *Celekli Abuzer. Determination of trophic weight and indicator values of diatoms in Turkish running waters for water quality assessment." Turkish Journal of Botany. (2019) 43(1): 90-101. https://doi.org/10.3906/bot-1704-40*

[9] *Ernazarovna Makhmudova Dildora, Buriev Eshmurod Sattorovich. Assessment Of Water Quality Of Small Rivers Of The Syrdarya Basins For The Safe Water Use. PalArch's Journal of Archaeology of Egypt/Egyptology. (2020) 17(7): 9901-9910.*

[10] *Solangi Ghulam Shabir. Groundwater quality evaluation using the water quality index (WQI), the synthetic pollution index (SPI), and geospatial tools: a case study of Sujawal district, Pakistan. Human and Ecological Risk Assessment: An International Journal. (2020) 26(6): 1529-1549. https://doi.org/10.1080/10807039.2019.1588099*

[11] *Camara Moriken, Nor Rohaizah Jamil, Ahmad Fikri Bin Abdullah. Impact of land uses on water quality in Malaysia: a review. Ecological Processes. (2019) 8(1): 1-10. https://doi.org/10.1186/s13717-019-0164-x*

[12] *Bisht Anil Kumar. Artificial neural network based water quality forecasting model for ganga river. International Journal of Engineering and Advanced Technology. (2019) 8(6): 2778-2785. https://doi.org/10.35940/ijeat.F8841.088619*

[13] *Hancock John T., Taghi M. Khoshgoftaar. Survey on categorical data for neural networks. Journal of Big Data. (2020) 7(1): 1-41. https://doi.org/10.1186/s40537-020-00305-w*

[14] *Kuha Jouni, Colin Mills. On group comparisons with logistic regression models. Sociological Methods & Research. (2020) 49(2): 498-525. https://doi.org/10.1177/0049124117747306*

[15] *Weizhuo Gai, Zhenyan Deng. A comprehensive review of adsorbents for fluoride removal from water: Performance, water quality assessment and mechanism. Environmental Science: Water Research & Technology. (2020) 7(8): 1362-1386. https://doi.org/10.1039/D1EW00232E*