Scholar Publishing Group
Water Pollution Prevention and Control Project
https://doi.org/10.38007/WPPCP.2020.010406
ISSN 2790-2730 Vol. 1, Issue 4: 51-60

Construction of Industrial Sewage Early Warning System Based on Bayesian Algorithm and Artificial Neural Network

Xiangru Hou*

Department of Information Engineering, Heilongjiang International University, Harbin 150025, China

houxiangru@hiu.net.cn

**corresponding author*

Keywords: Industrial Sewage, Sewage Early Warning System, Bayesian Algorithm, Artificial Neural Network

Abstract: In the process of industrialization, water pollution accidents occur from time to time, resulting in major economic losses and negative impacts on the water environment. Statistical analysis shows that it is difficult to issue early warning quickly and effectively before the accident, leading to unavoidable water pollution accidents. In order to improve the environment and reduce water pollution accidents, governments and enterprises at all levels need to invest a lot of human resources. In this paper, Bayesian algorithm and artificial neural network are used to study industrial wastewater early warning and effectively prevent water pollution accidents. This paper first introduces the main parts and functions of industrial sewage early warning, then uses Bayesian network to carry out water quality early warning, introduces the calculation process of Bayesian water quality evaluation, and establishes Bayesian network water quality early warning system. Then this paper uses artificial neural network to carry out water quality early warning, builds a comprehensive model of neural network evaluation, and introduces the process of artificial neural network water quality evaluation in detail. In the experimental part, Bayesian algorithm and artificial neural network algorithm are used to evaluate the grade of industrial sewage. The experimental results show that the two algorithms have good accuracy, F1 value and recall rate when evaluating the quality of industrial sewage. The accuracy rate of the industrial sewage quality evaluation results of the two algorithms is more than 95%, which shows that Bayesian algorithm and artificial neural network can be well applied in the industrial sewage early warning field.

1. Introduction

In recent years, due to urban expansion, population growth and rapid economic development, water pollution caused by industrial production activities has increased significantly. The waste produced by industrial production contains a large amount of waste water, and the water quality composition is complex. The concentration of lead and zinc is usually very high. The early warning system can monitor the wastewater containing cadmium ions and other toxic heavy metals that are harmful to the environment online, determine the source of pollutants, the concentration and type of pollutants discharged, and use monitoring tools as the basis for monitoring the pollutant discharge source, water quality early warning and environmental impact assessment.

At present, many scholars have studied industrial sewage. Guo Zhiwei uses image processing technology to study the decision-making process of complex sewage treatment [1]. Yang Jianqiao studied the supercritical water oxidation system for industrial sewage [2]. Fu Biao studied the thermochemistry and emission behavior during the combustion of industrial coal slurry and sewage sludge [3]. Bertocci I studied the relationship between sewage discharge in industrial areas and the community of small benthos [4]. Liu Xin-Cong carried out path analysis on chemical oxygen demand and ammonia nitrogen emission of industrial sewage [5]. Sarvari Hadi analyzed the obstacles of private sector investment in water supply and sewage treatment industry [6]. Although there are many studies on industrial sewage, how to carry out industrial sewage early warning is still a problem to be studied.

At present, Bayesian algorithm and artificial neural network have good early warning effect. Li Xiang used Bayesian network for earthquake parameter early warning [7]. Xie Xiaoliang used Bayesian network reasoning to monitor and early warning air pollution risks [8]. Aghaali Mohammad studied the role of Bayesian algorithm in the early warning of infectious disease outbreak [9]. Geng Zhiqiang studied the role of short-term and short-term memory neural network in food safety risk early warning [10]. Anbarasan M uses convolution depth neural network to study flood disaster early warning system [11]. Yang Beibei used short-term and short-term memory neural network to predict landslide displacement [12]. Although Bayesian algorithm and artificial neural network have been applied in various fields of early warning, the research on the application of Bayesian algorithm and artificial neural network in industrial sewage early warning is not comprehensive enough.

In order to improve the efficiency of industrial sewage early warning and better protect the environment, this paper uses Bayesian algorithm and artificial neural network to build the industrial sewage early warning system. First of all, the process of industrial wastewater treatment is introduced in detail, that is, determining the importance of early warning, determining the source of early warning, analyzing early warning signals, reporting the degree of prediction, and eliminating the hidden dangers of early warning; Then the Bayesian network water quality early warning system and the artificial neural network water quality early warning system are constructed, and the algorithms of the two systems are introduced in detail; Finally, Bayesian algorithm and artificial neural network are used to evaluate the industrial sewage grade, and the accuracy, F1 value and recall rate of the grade evaluation results are calculated to verify the effectiveness of the industrial sewage early warning system.

2. Industrial Sewage Early Warning

Industrial water pollution early warning is to apply the early warning principle to industrial water pollution control. The logical process of industrial water pollution early warning mechanism should include the following five steps: determining the importance of early warning, determining the source of early warning, analyzing early warning signals, reporting the degree of prediction, and eliminating early warning hidden dangers. The warning principle is shown in Figure 1.

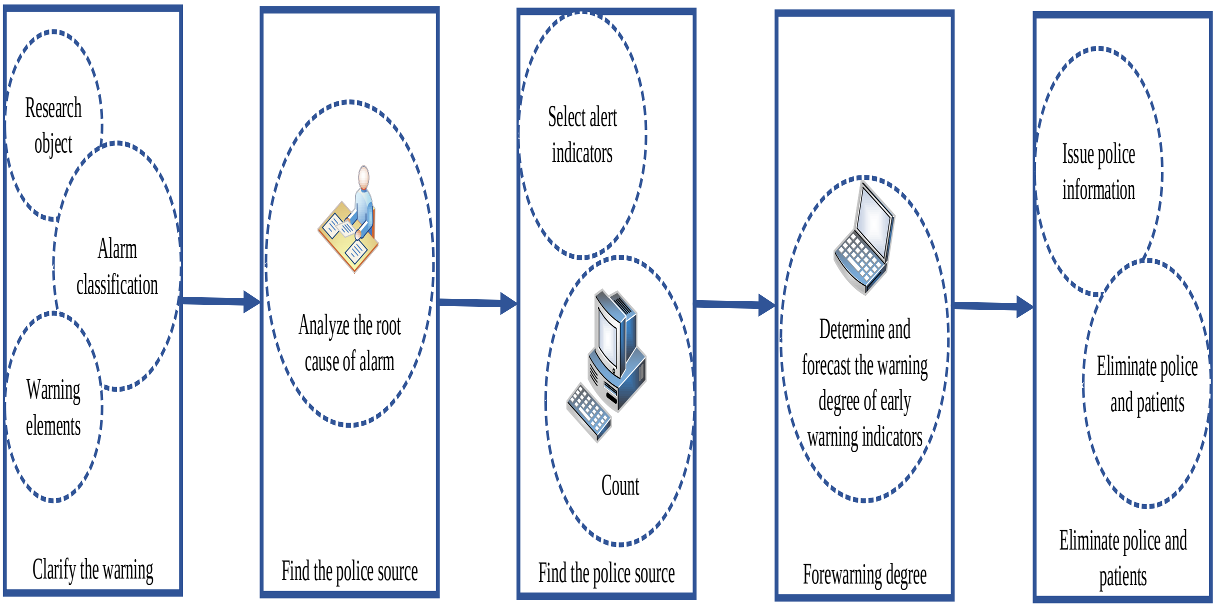


Figure 1. Industrial water pollution early warning principle

(1) Clarify the warning

It is clear that the warning meaning includes the elements and degree of early warning. Early warning elements are indicators of early warning conditions, which means the type of conditions that cause alarm in industrial waters. Industrial water warning refers to the prediction of dangerous points or dangerous areas. This process involves changing various factors affecting the quality of industrial water, and spreading the warning of industrial water management, providing a basis for monitoring and decision-making. These dangerous points or areas are actually very abnormal conditions in the development of industrial water system, and are called "early warning elements" in early warning science. The severity of early warning factors, that is, the danger degree of dangerous points or dangerous areas, is called early warning degree.

Decompose and improve the common factors of industrial water pollution early warning. According to the space-time dynamic analysis of regional industrial water quality and regional industrial water pollution risk assessment, the regional industrial water environmental pollution early warning level is divided into five levels [13].

(2) Find the source of warning

The warning source is the warning source, that is, the existing or potential pollution factors in industrial water. Warning sources can be divided into internal and external. The early warning sources generated under the natural background are objective information about natural abnormal changes, which may lead to early warning of natural disasters and industrial water pollution. In low humidity areas, the flow of industrial water is slow, various mobile elements accumulate, and the mineralization rate of industrial water is much higher than the default value. It is a warning source under natural background conditions. This warning source depends on internal factors, and the monitoring warning source is weak. The external source is the external warning source, and the human pollution source is the main warning source.

(3) Early warning signal analysis

Before warning, there is always a special warning sign. For example, the baseline value of industrial water quality affected by endogenous alarm sources, the natural conditions of aquifer protection, and the purification capacity of aquifer. However, in the development of industrial production, the distribution and change of pollution point sources, the distribution of pollution sources, the speed of wastewater treatment, and the speed of wastewater discharge that meet the requirements are not direct indicators of the prosperity of industrial water supply system, but trend indicators. Early warning signal analysis is the most critical and complex part in the early warning process. The key of early warning signal analysis is to identify the early warning signal indicators, and then further analyze the quantitative relationship between early warning signals to obtain various scenarios of early warning signals, and then predict the warning level of warning elements in the warning area of warning signs.

(4) Forewarning degree

The alert level is the target of the alert. For different alert conditions and alert indicators, the alert level predicts the scope of the system to analyze and report the "alert". There are usually two methods to predict the alert level. One of them is to create a common alarm unit model, and then switch to the alarm level according to the calculated alarm threshold. The second alarm level model related to alarm components is hierarchical regression technology. Alarms are generally divided into five levels, namely, no alarm, light alarm, medium alarm, heavy alarm and large alarm. These five levels correspond to the quantitative range of the index change of alarm elements, namely, the alarm limit. The key to determining the alarm level is to determine the limits of the alarm.

(5) Eliminate police and patients

In order to reduce or prevent the risk of industrial water pollution and the deterioration of industrial water quality, corresponding solutions are proposed.

3. Bayesian Network Water Quality Early Warning

3.1. Bayesian Network Water Quality Assessments

The assessment or prediction of water quality is actually an inference of water quality. That is to say, use all known information to infer different probabilities statistically and make the final prediction according to the maximum probability principle [14]. On this basis, the Bayesian network analysis method is determined:

$$P(B_i|A) = \frac{P(B_i)P(A|B_i)}{\sum_{i=1}^{S} P(B_i)P(A|B_i)} \tag{1}$$

B_i represents the water quality level, and A represents the sample water quality index value.

In this work, the formula (1) based on water quality assessment characteristics and water pollution data is defined as follows: the water quality grade value represented by B_i is expressed as y_{ij}. The water quality index of the observation station is expressed as x_j. I is the number of water quality grades, and j is the number of selected water quality indicators. Water quality is divided into five categories: I, II, III, IV and V, which can be defined as S=5. Formula (1) can be expressed as:

$$P(y_{ij}|x_j) = \frac{P(y_{ij})P(x_j|y_{ij})}{\sum_{i=1}^{S} P(y_{ij})P(x_j|y_{ij})} \tag{2}$$

Calculate $P(y_{ij})$. In many applications, if there is no prior information, it is difficult to determine which level the water quality of K station belongs to, but this can be determined in the water quality assessment. The most widely accepted principle is that if there is no water quality information, the possibility of water quality degradation is the same.

$$P(y_{1j}) = P(y_{2j}) = P(y_{3j}) = P(y_{4j}) = P(y_{5j}) = \frac{1}{5} \tag{3}$$

Calculate $P(x_j|y_{ij})$. A reasonable definition is the most important step in calculating Bayesian water quality estimation. The most commonly used distance method is selected based on intuitive, simple and principled reasons. According to the concept of geometric probability, the distance method is the distance between the water pollution source value and the default value of the water quality index.

$$P(x_j|y_{ij}) = \frac{1/L_{ij}}{\sum_{i=1}^{5} 1/L_{ij}} \quad (4)$$

$$L_{ij} = |x_j - y_{ij}| \quad (5)$$

Calculate $P(y_{ij}|x_j)$. Its significance is the probability that the observed value of j index is related to the Class I water at the known observation point.

Calculate the possibility of several indicators of the observation station to test the water quality as Grade I.

$$P_i = \sum_{j=1}^{m} w_j P(y_{ij}|x_j) \quad (6)$$

$$w_j = \frac{x_j/y_{ij}}{\sum_{j=1}^{m} x_j/y_{ij}} \quad (7)$$

Level f is derived from the maximum probability principle.

$$P_f = \max_{i=1-5} P_i \quad (8)$$

According to this procedure, water quality assessment can be carried out at the observation station.

3.2 Bayesian Network Water Quality Early Warning Systems

The Bayesian water quality assessment and prediction framework is shown in Figure 2.

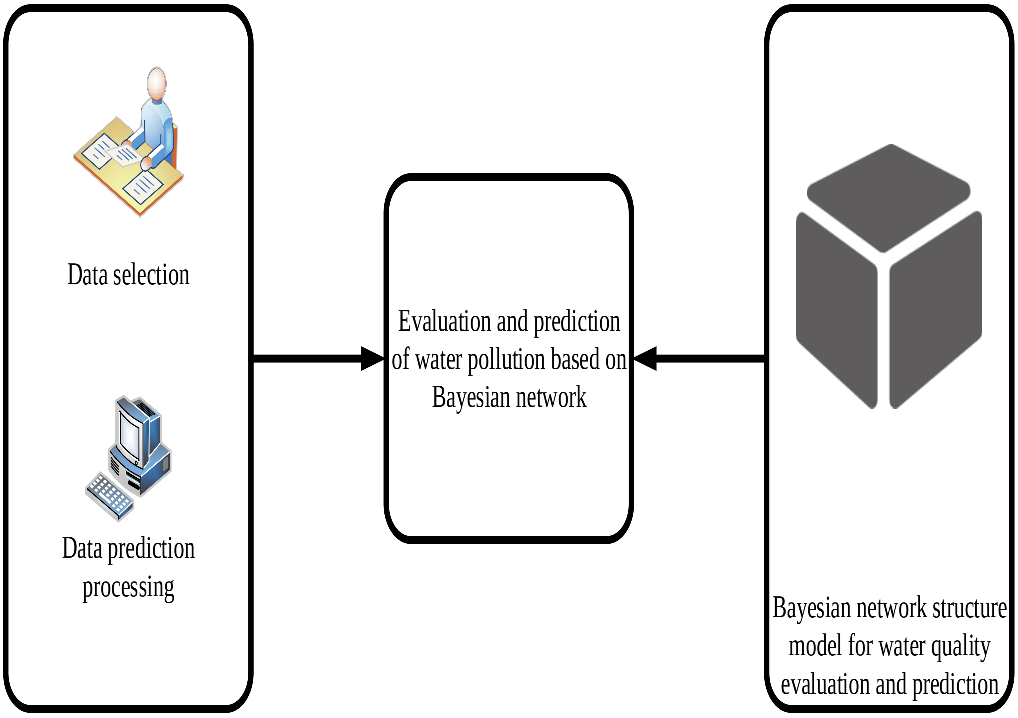


Figure 2. Framework system of water quality pollution assessment and prediction based on Bayesian network

(1) Data selection and forecast data processing. Data selection refers to the selection of specific sampling data and experimental data from the obtained monitoring data, which is mainly used to construct the Bayesian water quality assessment network structure. Prediction data processing is a group of processing data used for water quality prediction, mainly using experimental data, and using some data processing methods to obtain unknown data of future water pollution.

(2) Establishment of network structure. Use sampling data and specific data processing methods to determine the degree of correlation between network nodes.

(3) Water quality assessment and prediction. First, according to the different network configurations, the relationship between the node and the parent node to be solved can be detected (here is the water quality level and related pollution factors). Secondly, the calculation formula of water quality assessment and prediction is deduced from the relationship between nodes. Finally, the assessment or prediction of water resources is based on the prediction of water quality.

4. Artificial Neural Network Water Quality Early Warning

The input and expected output are A_K and C_K, and the number of input template pairs in the learning matrix is K. The initial connection weight from the input plane to the network hidden layer is w_{ij}, and the threshold value of the hidden layer and the output layer is θ_{i,γ_i}. The expected inputs and outputs are calculated as follows

Calculate the new value of hidden layer activation:

$$b_i = f(\sum_{h=1}^{n} a_h u_{hi} + \theta_i) \tag{9}$$

Activation value of output unit of calculation layer:

$$c_j = f(\sum w_{ij} b_i + y_i) \tag{10}$$

Calculation error of output layer:

$$d_j = c_j(1 - c_j)(c_j^k - c_j) \tag{11}$$

Calculate the error of each d_j element hidden layer unit.

$$e_i = b_i(1 - b_i) \sum_{j=1}^{q} w_{ij} d_j \tag{12}$$

Change output layer unit threshold:

$$\Delta r_i = a d_j \tag{14}$$

Manage connections between imported and hidden layer cells:

$$\Delta V_{hi} = \beta a_h e_i \tag{15}$$

Adjust hidden layer unit threshold:

$$\Delta V_{hi} = \beta a_h e_i \tag{15}$$

$$\Delta \theta_i = \beta e_i \tag{16}$$

These calculation steps are repeated until the output error of all samples is less than a certain convergence error, and the learning is completed. Based on these defined connection weights and thresholds, a comprehensive model of neural network evaluation can be obtained [15].

The process of evaluating water quality using BP neural network (Back Propagation) is shown in Figure 3.

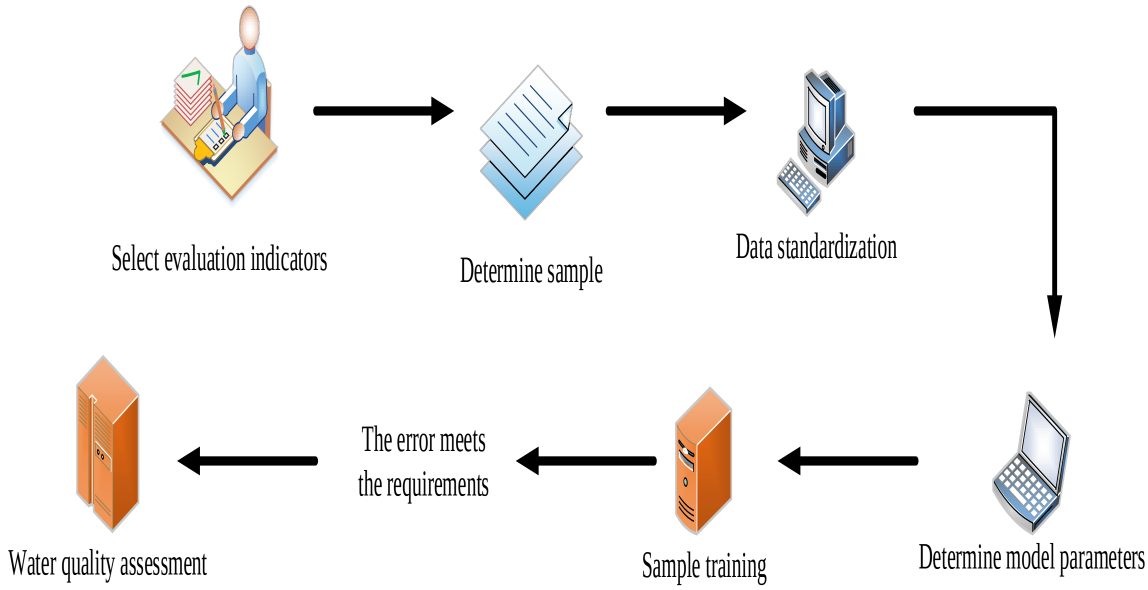


Figure 3. BP neural network water quality evaluation

First of all, the evaluation indicators are selected according to the specific conditions of the surveyed area and the reference points in the water quality monitoring data. Second, determine the total number of sample drivers and the occurrence value of each input parameter in each sample, and standardize all input and output data. Third, determine the number of nodes in the hidden layer, the number of nodes in the output layer, the learning efficiency, the retention error and other model parameters, and determine the original model of the neural network. Fourth, input the learning mode and check whether it is within the allowable error range.

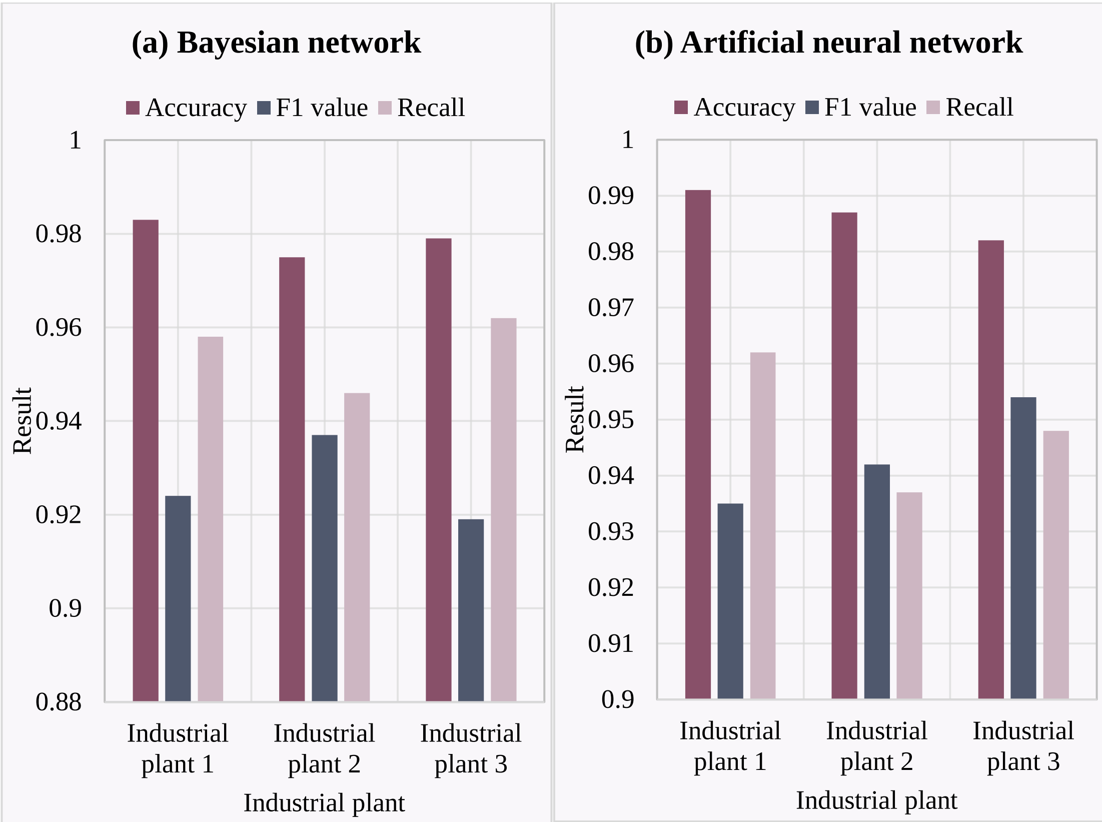
The standardization of raw data is necessary to eliminate the impact of different measurements and values on the calculation process caused by the differences in the characteristics of different monitoring indicators.

$$x_i^{k*} = \frac{x_i^k}{x_{i,max} - x_{i,min}} \tag{17}$$

5. Industrial Sewage Quality Early Warning Experiment

The warning standard in this paper is the national industrial sewage discharge standard. Bayesian network and artificial neural network are used to predict the water quality of industrial sewage in three factories, and the accuracy, F1 value and recall rate of the prediction results of the industrial sewage water quality evaluation results are calculated.

The evaluation results of industrial sewage quality are shown in Figure 4.



(a) Evaluation results of industrial wastewater quality based on Bayesian network
(b) Evaluation results of industrial wastewater quality based on artificial neural network

Figure 4. Evaluation results of industrial sewage quality

Figure 4 (a) shows the evaluation results of industrial wastewater quality by Bayesian network, and Figure 4 (b) shows the evaluation results of industrial wastewater quality by artificial neural network.

When Bayesian network is used to evaluate the industrial sewage quality of plant 1, the accuracy rate of the water quality evaluation result is 0.983, the F1 value of the water quality evaluation result is 0.924, and the recall rate of the water quality evaluation result is 0.958. When Bayesian network is used to evaluate the industrial sewage quality of plant 2, the accuracy rate of the water quality evaluation result is 0.975, the F1 value of the water quality evaluation result is 0.937, and the recall rate of the water quality evaluation result is 0.946. When Bayesian network is used to evaluate the industrial wastewater quality of plant 3, the accuracy rate of the water quality evaluation result is 0.979, the F1 value of the water quality evaluation result is 0.919, and the recall rate of the water quality evaluation result is 0.962.

When the artificial neural network is used to evaluate the industrial wastewater quality of industrial plant 1, the accuracy rate of the water quality evaluation result is 0.991, the F1 value of the water quality evaluation result is 0.935, and the recall rate of the water quality evaluation result is 0.962. When the artificial neural network is used to evaluate the industrial wastewater quality of industrial plant 2, the accuracy rate of the water quality evaluation result is 0.987, the F1 value of the water quality evaluation result is 0.942, and the recall rate of the water quality evaluation result is 0.937. When the artificial neural network is used to evaluate the industrial wastewater quality of industrial plant 3, the accuracy rate of the water quality evaluation result is 0.982, the F1 value of the water quality evaluation result is 0.954, and the recall rate of the water quality evaluation result is 0.948.

When Bayesian network and artificial neural network are used to evaluate the quality of industrial sewage, the accuracy rate of the evaluation results of industrial sewage quality is above 0.95, which shows that both algorithms have good water quality evaluation ability and can monitor and early warning the water quality.

6. Conclusion

In this paper, Bayesian algorithm and artificial neural network are used to study the industrial sewage early warning system. This paper first introduces the process of industrial sewage early warning, then constructs the Bayesian water quality early warning system according to the principle of Bayesian network, and then constructs a comprehensive model of neural network evaluation according to the principle of artificial neural network. In the experiment part, Bayesian algorithm and artificial neural network are used to predict the quality of industrial sewage from three factories, and the accuracy, F1 value and recall rate of the prediction results are calculated. The results show that the two water quality early warning systems have good water quality prediction effect and can be better applied in the field of industrial sewage early warning.

Funding

This article is not supported by any foundation.

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] Guo, Zhiwei. "Graph embedding‐based intelligent industrial decision for complex sewage treatment processes." International Journal of Intelligent Systems 37.12 (2020): 10423-10441.

[2] Yang, Jianqiao. "Novel design concept for a commercial-scale plant for supercritical water oxidation of industrial and sewage sludge." Journal of environmental management 233.3 (2019): 131-140.

[3] Fu, Biao. "Co-combustion of industrial coal slurry and sewage sludge: Thermochemical and emission behavior of heavy metals." Chemosphere 233.10 (2019): 440-451.

[4] Bertocci, I. "Multiple human pressures in coastal habitats: variation of meiofaunal assemblages associated with sewage discharge in a post-industrial area." Science of the Total Environment 655.3 (2019): 1218-1231.

[5] Liu, Xin-Cong. "A path analysis for chemical oxygen demand and ammonia nitrogen discharge from industrial sewage in China." Water Resources 47.1 (2020): 1012-1019.

[6] Sarvari, Hadi. "Barriers to development of private sector investment in water and sewage industry." Built Environment Project and Asset Management 11.1 (2020): 52-70.

[7] Li, Xiang. "Dynamic early warning of rockburst using microseismic multi-parameters based on Bayesian network." Engineering Science and Technology, an International Journal 24.3 (2020): 715-727.

[8] Xie, Xiaoliang. "Bayesian network reasoning and machine learning with multiple data features: air pollution risk monitoring and early warning." Natural Hazards 107,1 (2020): 2555-2572.

[9] Aghaali, Mohammad. "Performance of Bayesian outbreak detection algorithm in the syndromic surveillance of influenza - like illness in small region." Transboundary and Emerging Diseases 67.5 (2020): 2183-2189.

[10] Geng, Zhiqiang. "Risk early warning of food safety using novel long short-term memory neural network integrating sum product based analytic hierarchy process." British Food Journal 124.3 (2020): 898-914.

[11] Anbarasan, M. "Detection of flood disaster system based on IoT, big data and convolutional deep neural network." Computer Communications 150.1 (2020): 150-157.

[12] Yang, Beibei. "Time series analysis and long short-term memory neural network to predict landslide displacement." Landslides 16.1 (2019): 677-694.

[13] Panchal, Deepak. "SARS-CoV-2: sewage surveillance as an early warning system and challenges in developing countries." Environmental Science and Pollution Research 28 .4(2020): 22221-22240.

[14] Wepener, V., and G. C. O'Brien. "The application of Bayesian networks to evaluate risks from multiple stressors to water quality of freshwater ecosystems." African Journal of Aquatic Science 47.3 (2020): 231-244.

[15] Kadam, A. K. "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 5.3 (2019): 951-962.