UDC 004.8

INTELLIGENT ALGORITHMS TO ASSIST IN MAKING WATER QUALITY DECISIONS BASED ON ARTIFICIAL INTELLIGENCE MODELS

[+]*Nazarov F.*[1], *Salimova M.*[1], *Shukurov B.*[1]

[1] Samarkand State University named after Sharof Rashidov, Samarkand, Uzbekistan

[+] maxbubacoc@gmail.com

Abstract. This article examines methods of analyzing water composition and the possibilities of using modern technologies. The study compared traditional physical, chemical, and biological methods of water quality assessment, as well as measurements performed using sensor technologies. The possibility of quickly and accurately determining water parameters using devices such as the Exo2 sonde and TDS sensor was considered. In addition, it was shown that artificial intelligence-based algorithms can classify water samples in water quality assessment and analysis, evaluate them according to different parameters, and group them into categories. The application of Decision Tree, Random Forest, and SVM models was also analyzed. The results of the study examined the possibilities of using water monitoring systems as an effective tool that increases efficiency and accuracy.

Keywords: WQI, sensor technology, artificial intelligence, machine learning, decision tree, random forest, SVM, min-max scaling, SMOTE algorithm.

1 INTRODUCTION

Water is considered the source of life for humans and all living organisms, and only 2.5% of the total water on Earth is regarded as suitable for drinking [1]. The growing world population and improving living standards are increasing the demand for water. In particular, over the past 100 years, global demand for freshwater has increased by 600% and continues to grow by approximately 1.8% annually [2]. In Uzbekistan, an average of 4.7 thousand liters of water is consumed per capita per day, and a total of 54.5 billion liters of water are consumed annually across the country.

To meet the needs of living organisms, certain requirements regarding water quality and condition must be fulfilled and maintained. Two systems have been developed and are used for water quality assessment: Water Quality Index (WQI) and Interim National Water Quality Standards (INWQS). The WQI helps to evaluate the condition and overall status of water, while INWQS is used to classify how water bodies can be utilized based on WQI data. According to INWQS, the quality of water bodies is divided into six classes (Class I, Class II, Class II, Class III, Class IV, and Class V) [3]. Here, Class I represents the best water quality condition, while Class V indicates the worst condition, where extensive treatment is required for practical use.

2 RELATED WORK

The detection and analysis of pollutants in water is a complex process, and there are several methods of analysis.

Physical method is a simple and rapid method for assessing the overall condition of water. Physical analysis of water includes the evaluation of various properties such as temperature, color, turbidity, and conductivity. These parameters provide valuable information about the clarity and suitability of water for different purposes.

Turbidity indicates the presence of particles such as sediments, algae, or organic matter. High turbidity can affect the aesthetic quality of water and cause problems in disinfection processes. Color reflects the presence of suspended or dissolved substances in water and affects its aesthetic and technological properties. Temperature plays an important role in physical and biological processes in aquatic environments, particularly influencing dissolved oxygen levels and gas solubility. Conductivity measures the ability of water to conduct electric current and provides information about the total mineral content and salinity of water, as well as its suitability for various applications and industrial use.

Chemical analysis is a fundamental part of water quality assessment. It requires testing various chemical parameters to identify pollutants and evaluate their levels in water (Figure 1). Commonly tested chemical parameters include ammonia, chloride ion, nitrite, nitrate, phosphate, and water hardness [4].

Chlorite (ClO_2^-) is formed as a result of disinfection with chlorine dioxide. It must be present in limited amounts in drinking water because it can affect the blood system and nervous system. Nitrate (NO_3^-) mainly enters water through fertilizers used in agriculture. When consumed in high amounts, it increases the risk of methemoglobinemia, especially in infants. Nitrite (NO_2^-) is more toxic than nitrate and disrupts oxygen transport in the blood. It must be present in very low amounts in drinking water. Ammonia (NH_3) is commonly found in water sources. High levels of ammonia in drinking water have a negative effect on health. Phosphate (PO_4^{3-}) mainly originates from agricultural waste, sewage, and industrial effluents. Although phosphate itself is not toxic, high concentrations can lead to eutrophication of water. Water hardness is determined by the presence of calcium (Ca^{2+}) and magnesium (Mg^{2+}) salts in water. Hard water can be beneficial to some extent for drinking, but excessive amounts of certain ions cause inconvenience due to scale formation [5].

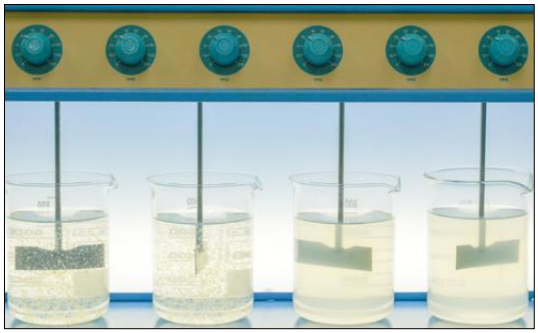

Fig. 1. Process of chemical analysis of water

Biological analysis is aimed at assessing the presence of microorganisms and other biological indicators in water. Drinking water should be completely free of microbiological indicators such as Intestinal enterococci and E. coli [6].

Bacterial testing is important in determining the safety and suitability of water for drinking. One of the most common bacteria found in water is E. coli, and its high concentration leads to contamination of water with pathogenic microorganisms that are considered dangerous to human health (Figure 2).

Physical, chemical and biological methods are traditional methods for determining the quality and composition of water. These methods provide scientifically accurate and reliable results, but are often time-consuming, require laboratory conditions and qualified specialists. Therefore, real-time monitoring and rapid analysis of large volumes of data are difficult.

In recent years, sensor technologies have opened up new opportunities for monitoring water quality. Sensors automatically measure temperature, pH, electrochemical indicators, TDS, TSS and other parameters in real time. This significantly simplifies the process of monitoring water quality and increases accuracy. Such devices are designed to measure individual water parameters, but based on multifunctional sensor technologies, they allow for the analysis of several parameters simultaneously in real time.

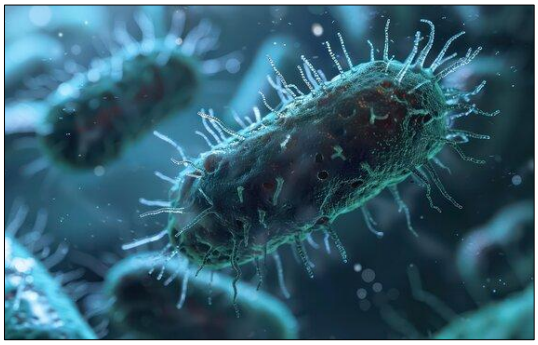

Fig. 2. E.coli bacteria

The EXO2s device is a modern monitoring device designed to automatically measure the physical, chemical and biological parameters of water. This device allows for long-term, continuous and unattended water quality monitoring (Figure 3 (a)).

The probe has a robust design and offers the possibility of integration with various monitoring systems. The device stands out from other conventional water quality measuring devices with its high-precision data acquisition functions even in extreme environmental conditions. The EXO2s model has 7 universal sensor ports, through which it is possible to measure several parameters simultaneously (Figure 3 (b)). The device

is also corrosion-resistant and equipped with a central anti-pollution automatic cleaning system, which ensures long-term efficient operation of the sensors.

a) b)

Fig. 3. a) EXO2s sonde; b) Internal structure of the EXO2s probe

The device is powered by an external power source and has a large internal memory, allowing for long-term storage of measurement results. The EXO2s probe is automatically calibrated, has a battery life of over 3 months, and is designed for use at depths of up to 250 meters [7]. It also provides high-precision real-time measurements of water composition parameters. The water parameters measured by the device are listed in Table 1.

Table 1. Traditional methods of water quality analysis

Category	Parameters
Physical	Temperature, turbidity, depth, pressure, density
Chemical	pH, electrical conductivity, salinity, dissolved oxygen, nitrate, ammonium, ammonia, chloride, total dissolved solids
Biological	Chlorophyll content, algae content, organic matter content

The EXO2s multi-parameter probe is widely used in water quality monitoring in rivers and lakes, monitoring drinking water sources, studying sea and ocean waters, monitoring groundwater, monitoring wastewater, in aquaculture systems, and monitoring industrial water.

Total Dissolved Solids (TDS) sensors are important devices used to measure the concentration of dissolved solids in water (Figure 4). A TDS sensor indicates the total concentration of all dissolved solids in water. This includes salts, minerals, and organic matter. TDS is usually expressed in milligrams per liter (mg/L) or parts per million (ppm). High TDS levels can affect water quality, affecting the taste, color, and health standards of the water. The acceptable level of TDS depends on the purpose for which the water is used. For example, drinking water should generally have a TDS level below 500 mg/L, while aquaculture and hydroponic systems have special requirements for TDS [8].

TDS sensors are easy to use and can be used in both laboratory and field conditions. Regular calibration, temperature, and sensor cleanliness are important when using the sensor. TDS sensors are widely used in various industries (Table 2).

Fig. 4. TDS sensor

TDS sensors are an important tool in monitoring and maintaining water quality in various industries. Understanding how they work and correctly interpreting the data obtained helps develop effective water resource management strategies.

Table 2. TDS sensor applications

Field	Application
Water quality monitoring	Regular monitoring of urban water sources
Aquaculture	Maintaining optimal water quality for aquatic organisms; monitoring salinity levels
Hydroponics	Monitoring nutrient levels in nutrient solutions and adjusting fertilizer concentration
Industry	Monitoring water used in cooling at power plants and in production processes
Aquarium maintenance	Regulating the addition of new water based on TDS levels

3 MATERIALS AND METHODS

During the study, a step-by-step algorithm for the water quality assessment process was developed. Figure 5 shows a block diagram of the algorithm. Each block reflects the main stages of data processing and classification. The dataset analysis process was carried out according to this algorithm.

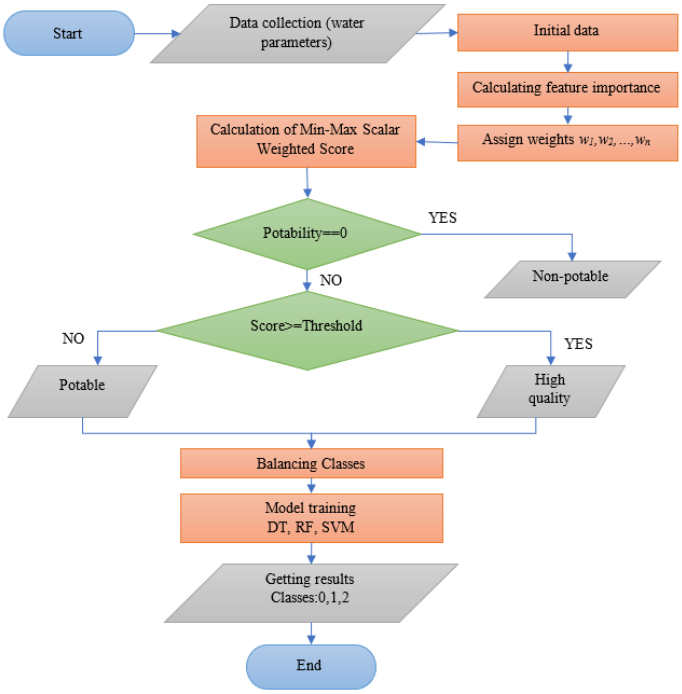

Fig. 5. Water quality classification algorithm

The dataset from kaggle.com was used to analyze water quality. The dataset consists of 10 features and 3276 objects in total. Figure 6 shows the structure of the dataset. In this study, a dataset consisting of physicochemical parameters was used to assess water quality. The dataset includes the following main parameters: pH, Hardness, TDS, Chloramines, Sulfate, Conductivity, Organic Carbon, Trihalomethanes, and Turbidity. Potability column, which represents the potability of water, was selected as the target variable.

	pH	Hardness	TDS	Chloramines	Sulfate	Conductivity	Organic_Carbon	Trihalomethanes	Turbidity	Potability
0	NaN	204.890455	20791.318981	7.300212	368.516441	564.308654	10.379783	86.990970	2.963135	0
1	3.716080	129.422921	18630.057858	6.635246	NaN	592.885359	15.180013	56.329076	4.500656	0
2	8.099124	224.236259	19909.541732	9.275884	NaN	418.606213	16.868637	66.420093	3.055934	0
3	8.316766	214.373394	22018.417441	8.059332	356.886136	363.266516	18.436524	100.341674	4.628771	0
4	9.092223	181.101509	17978.986339	6.546600	310.135738	398.410813	11.558279	31.997993	4.075075	0

Fig. 6. Dataset structure

Initially, only potable water samples (Potability = 1) were selected, and their features were normalized. During the normalization process, the Min-Max scaling method was mainly used to bring parameter values into the same range [9]. Using this method, in the next stage, the impact level of the physical and chemical

parameters of water on its potability was evaluated using the Random Forest Classifier algorithm. Through the model, the influence levels of each parameter in the overall classification process were determined and ranked according to their importance (Figure 7). The results showed that the parameters pH, Sulfate, and Hardness had a high impact and stood out as the main factors determining water quality. TDS and Chloramines had a moderate level of influence, while Conductivity and Organic Carbon were evaluated as auxiliary factors. Turbidity and Trihalomethanes were included among the parameters with relatively low impact. These results are also consistent with WHO standards, according to which pH is considered the parameter with the highest impact on water quality.

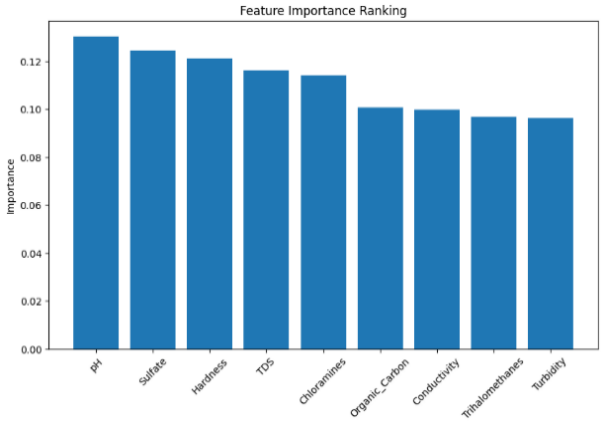

Fig. 7. Feature importances of input parameters

In the next stage, to more accurately assess water quality, weight coefficients were assigned to the parameters based on their importance levels. In particular, higher weights were assigned to important indicators such as pH, Sulfate, and Hardness. Based on these weights, an overall evaluation indicator (weighted score) was calculated for each sample. Based on the calculated values, an optimal threshold value (threshold = 5.35) was selected, and potable water was divided into two categories: potable (Status = 1) and high quality (Status = 2). Non-potable water (Potability = 0) was assigned as a separate class (Status = 0). As a result, the dataset was enriched with a new Status column consisting of three classes. Figure 8 presents the distribution of classified water quality based on the given data.

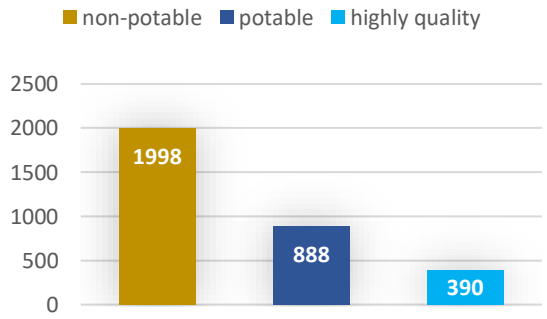

Fig. 8. Feature importances of input parameters

To eliminate class imbalance in the dataset, the SMOTE algorithm was used. This method enabled balancing the data by generating synthetic samples for minority classes. As a result, balance was achieved among all classes, and the efficiency of the model training process was improved.

4 ANALYSIS OF WATER COMPOSITION USING ARTIFICIAL INTELLIGENCE ALGORITHMS

In the process of water quality assessment, there is a need to analyze large-scale multidimensional data, which is limited in effectiveness when carried out using traditional methods. Therefore, the use of artificial intelligence algorithms has become widespread in modern research. These algorithms enable the analysis

of water composition, determination of quality indicators, and their automated classification. The main methods include regression, classification, clustering, and neural network models.

One of the classification algorithms of machine learning is the Decision Tree algorithm. A decision tree is a supervised learning method that can be used for both classification and regression problems, but it mainly shows good performance in solving classification problems. The decision tree uses a tree representation to solve problems. It is called a decision tree because, like a tree, it starts from a root and branches into subsequent nodes, forming a tree-like structure. In this case, decisions or tests are made based on the features of a given dataset. The decision tree produces a graphical representation to obtain all possible solutions to a problem/decision based on given conditions. The decision tree provides a condition and, based on the answer (Yes/No), splits into branches [10-11].

Decision tree is used for splitting and classification of data. The tree splitting criteria use the following main criteria for quality evaluation:

Gini index measures the quality of a split in the tree and indicates the probability of incorrectly classifying a randomly selected sample. This process can be seen in formula (1):

$$Gini(S) = 1 - \sum_{i=1}^{n} p_i^2, \tag{1}$$

here, S – the dataset, n – the number of classes, p_i^2 – the proportion of samples belonging to class i. The value ranges from 0 to 1, a lower value indicates a better split.

Information Gain (IG) measures the effectiveness of a split through entropy. This process can be seen in formulas (2) and (3):

$$E(S) = -\sum_{i=1}^{n} p_i \log_2(p_i), \tag{2}$$

$$IG(S, A) = E(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} E(S_v), \tag{3}$$

where, S_v – the subset of samples where attribute A has value v, $Values(A)$ – the set of all possible values of attribute A, p_i – the proportion of class i in the subset. Higher IG value indicates that the attribute is more effective for splitting.

Information Gain Ratio (IGR) is used to reduce the bias of multi-valued attributes in algorithms and is expressed by the following formula (4):

$$IGR(S, A) = \frac{IG(S, A)}{SplitInfo(S, A)}. \tag{4}$$

The Chi-square splitting criterion tests the relationship between an attribute and the class. This process can be seen in formulas (5) and (6):

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{k} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}, \tag{5}$$

$$E_{ij} = \frac{R_i \cdot C_j}{N}, \tag{6}$$

here, r – the number of categories of the attribute, k – the number of classes, O_{ij} – the observed frequency in the i-th category and j-th class, E_{ij} – the expected (unobserved) value, R_i – the total number of observations in the i-th category, E_j – the total number of observations in the j-th class, N – the total number of all observations.

Random Forest is an algorithm that works by combining multiple Decision Trees. It is used to solve classification and regression problems. The main idea of Random Forest is that several independent trees are constructed, and their results are combined using majority voting or averaging to produce the final decision [12]. This process is shown in Figure 9.

Preparation of training data: The given training dataset for the Random Forest algorithm is expressed as follows (7):

$$D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}. \tag{7}$$

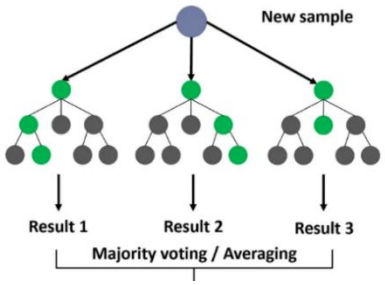

Fig. 9. Random forest model

Using the bootstrap method, B different subsets are generated. Based on each subset, a separate decision tree $h_b x$ is constructed ($b = 1, 2, ..., B$).

Tree construction: During the construction of each tree, a random subset of features is selected from all features for splitting. For example, if the total number of features is, then at each split, a random subset of features is selected.

Classification: In the Random Forest algorithm, the results of all trees are combined, and the final decision is determined based on majority voting [13-14]. This process can be seen in formula (8):

$$\hat{y} = \text{mode}\{h_1(x), h_2(x), \ldots, h_B(x)\}, \tag{8}$$

here, $\hat{y}$ – the class predicted by the model, $h_b x$ – represents the most frequently occurring value, B – the number of trees, mode – the function that selects the most frequent value.

If the problem is regression, the average of the results of all trees is taken, and this is expressed by the following formula (9):

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} h_b(x). \tag{9}$$

Support Vector Machine (SVM) is a model widely used for solving classification and regression problems. This algorithm is effectively applied in water quality assessment for classification of samples. SVM separates classes by mapping input data into a high-dimensional space and determining the maximum margin between them. The main goal of SVM is to separate two classes in the feature space and maximize the margin between the classes [15]. For this purpose, a separating hyperplane is constructed. SVM is mainly used in pattern recognition problems. For example, Figure 10 illustrates SVM classification with two classes that can be linearly separated by a hyperplane. The points closest to the hyperplane are called support vectors.

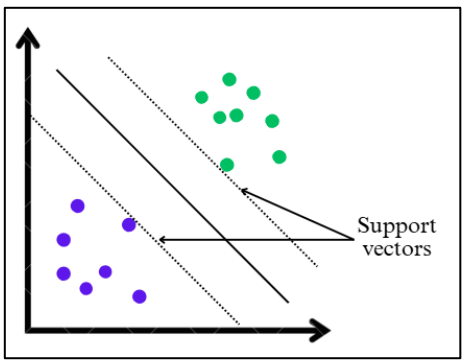

Fig. 10. SVM model

One of the important aspects of this algorithm is the use of kernel functions, which allow data to be mapped into a higher-dimensional space and enable effective classification of non-linearly separable data. In practice, several main kernel functions are widely used in the SVM algorithm [16-17].

Linear kernel is suitable for linearly separable data, works based on a simple scalar product, and is one of the simplest and fastest computational methods. This kernel function is expressed by formula (10):

$$K(x, y) = x^T y, \tag{10}$$

here, x and y are input data vectors, and x^T denotes the transposed form of the vector x.

Polynomial kernel enables the identification of more complex relationships by mapping data into a higher-dimensional space. This kernel is especially effective in cases where non-linear relationships exist. This kernel function is defined by the following formula (11):

$$K(x, y) = (x^T y + 1)^d, \tag{11}$$

here, d is the degree of the polynomial, which determines the complexity of the model. As the degree increases, the model becomes capable of learning more complex boundaries; however, the risk of overfitting also increases.

Radial Basis Function kernel is one of the most commonly used kernels. This kernel is based on the distance between data points and is highly effective in identifying complex, non-linear structures. This process can be seen in formula (12):

$$K(x, y) = exp(-\gamma \| x - y \|^2), \tag{12}$$

here, $\| x - y \|^2$ is the squared Euclidean distance between two vectors, and γ is the parameter that defines the influence radius of the kernel. If the value of γ is large, the model focuses more on local features, if it is small, it learns more general relationships.

Sigmoid kernel works similarly to activation functions in neural networks, and in some cases this kernel produces results similar to artificial neural networks. The Sigmoid kernel function is mathematically expressed by formula (13):

$$K(x, y) = tanh(\alpha x^T y + c), \tag{13}$$

here, α is the scalar coefficient, and c is the bias coefficient; these determine the flexibility of the model. This kernel provides behavior similar to neural networks.

The choice of kernel function depends on the type of problem, the characteristics of the data, and their distribution. In practical studies, the RBF kernel is widely used because it provides high accuracy. At the same time, proper tuning of kernel parameters (e.g., γ, d, α) is important to ensure optimal model performance [18-19].

By extending the SVM method to regression and prediction tasks, the SVR model is expressed by formula (14):

$$f(x) = \sum_{i=1}^{n} \alpha_i K(x, x_i) + b, \tag{14}$$

here, x – input data, n – number of input samples, α and b – Lagrange multiplier and bias term, $K()$ – kernel function.

SVM algorithm is one of the effective machine learning methods that enables efficient separation of classes in a high-dimensional space, identification of complex non-linear relationships, and solving classification and regression problems with high accuracy.

5 EXPERIMENTAL RESULTS

In the study, Decision Tree, Random Forest, and SVM models were sequentially applied for water quality classification. The model accuracy results are presented in Table 3. Initially, the Decision Tree algorithm was used as a baseline model, and it showed an accuracy of 0.68.

In the next stage, the Random Forest model was applied. This model achieved higher accuracy across all classes, reaching an overall accuracy of 0.81. In particular, it demonstrated superior performance compared to other models in identifying non-potable water (class 0).

The SVM model was also applied, showing an accuracy of 0.76. Although this model achieved high results in identifying class 1 and class 2, the error rate for class 0 remained relatively higher.

Table 3. Overall comparison of models

Model	Accuracy	Precision	Recall	F1-score
Decision Tree	0.68	0.68	0.68	0.68
Random Forest	0.81	0.81	0.81	0.80
Support Vector Machine	0.76	0.76	0.76	0.74

Figure 11 shows the confusion matrices for the Decision Tree, Random Forest, and SVM models. These matrices visually present the classification accuracy and errors of each model across the classes.

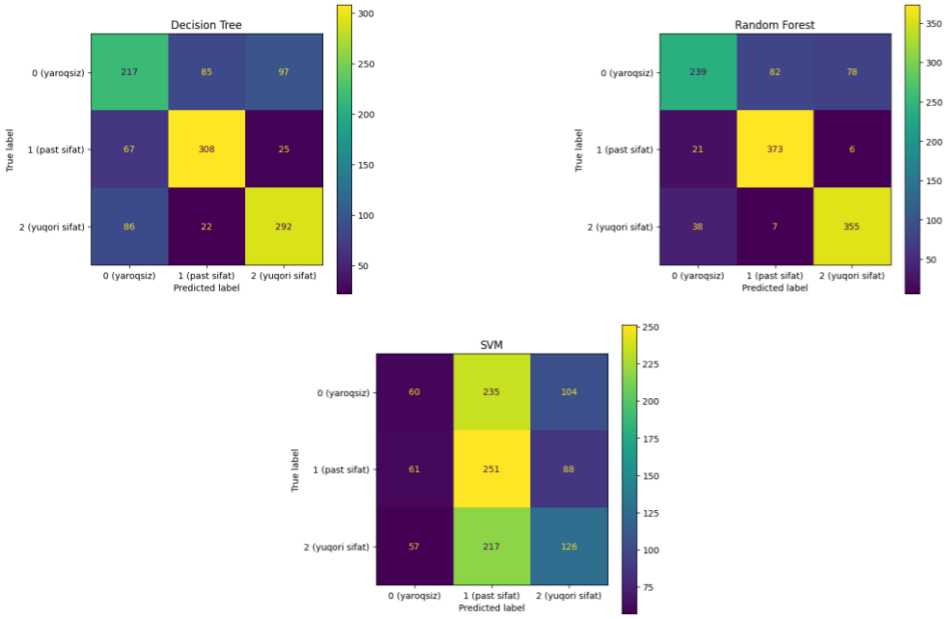

Fig. 11. Confusion matrices of the models

According to the overall analysis results, the Random Forest model was evaluated as the most effective model, providing the highest accuracy and the most balanced results.

6 CONCLUSION

In this study, traditional and modern methods of water composition analysis were investigated. The possibilities of water quality assessment using physical, chemical, and biological methods were analyzed, and the capabilities of devices such as the Exo2 sonde and TDS sensor for water quality monitoring were also examined. It was shown that artificial intelligence algorithms, including Random Forest, Decision Tree, and SVM models, provide opportunities for automated classification of water samples, evaluation based on different parameters, and effective decision-making. According to the obtained results, the Decision Tree model demonstrated an accuracy of 0.68 and served as a baseline approach. The Random Forest model achieved the highest performance, reaching 0.81 accuracy, 0.81 precision, 0.81 recall, and 0.80 F1-score. This model provided stable and balanced results across all classes. The SVM model showed an accuracy of 0.76, demonstrating particularly high performance in identifying class 1 and class 2, while the accuracy for class 0 (non-potable water) remained relatively lower. Based on the overall analysis, the Random Forest model was selected as the most effective and reliable model for water quality classification. This approach enables high-accuracy determination of water potability in real-world applications. The results of the study contribute to improving the efficiency of water monitoring systems and strengthening the classification-based analysis of water quality.

REFERENCES

[1] *Mishra R.K., Dubey S.C.* Fresh Water Availability and Its Global Challenge. Journal of Marine Science and Research, 2023.

[2] *Boretti A., Rosa L.* Reassessing the projections of the World Water Development Report. npj Clean Water, 2019. https://doi.org/10.1038/s41545-019-0039-9.

[3] *Misman N.A., Sharif M.F., Chowdhury A.J.K., Azizan N.H.* Water pollution and the assessment of water quality parameters: a review. Desalination and Water Treatment, 294, 79–88, 2023.

[4] Atlas Scientific. Water Analysis Methods. Available at: https://atlas-scientific.com/blog/water-analysis-methods/ (accessed: 10 April 2026).

[5] *Khamidov I.D., Akhatov A.A., Salimova M.Kh.* Determining the Potability of Water Using Machine Learning Based on Data Preprocessing. International Journal of Artificial Intelligence, 2026, Vol. 6, No. 3, pp. 959.

[6] European Commission. Drinking Water: Essential Quality Standards. Available at: https://eur-lex.europa.eu/EN/legal-content/summary/drinking-water-essential-quality-standards.html

[7] YSI (Xylem Inc.). EXO2s Multiparameter Sonde. 2024. Available at: https://www.ysi.com/exo2s (accessed: 10 April 2026).
[8] *Kumar V., Sharma A., et al.* Water Quality Prediction Using Machine Learning Algorithms: A Review. Journal of Environmental Management, 2022.
[9] *Nazarov Fayzulla, Akhatov Abror*. Methods for preparing a dataset for determining soil fertility based on artificial intelligence models . (2025). Digital transformation and artificial intelligence, 3(2), 91-99. https://dtai.tsue.uz/index.php/dtai/article/view/v3i215
[10] *Guido R., Ferrisi S., Lofaro D., Conforti D.* An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review. Information, 15(4), 235, 2024.
[11] *Nazarov F., Yarmatov Sh.* Sun'iy intellekt asoslari. 2025, pp. 135–143.
[12] *Yarmatov S., Rai H., Absalamova G., Absalamova D.* Real Estate Valuation based on XGBoost. In: Proceedings of the 8th International Conference on Future Networks & Distributed Systems (ICFNDS '24). ACM, 2025, pp. 705–709. https://doi.org/10.1145/3726122.3726223
[13] *Fayzullo N., Sariyev S., Sherzodjon Y.* Analyzing the Effectiveness of Ansambl Methods in Solving Multi-Class Classification Problems. 2025 International Russian Smart Industry Conference (SmartIndustryCon), Sochi, Russia, pp.788–793.
[14] *Dewi D.A., Wei A.S., Lin C.L., Heng C.D.* Water Quality Prediction using Random Forest Algorithm and Optimization. Journal of Applied Data Sciences, 2024.
[15] *Guido R., Ferrisi S., Lofaro D., Conforti D.* An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review. Information, 15(4), 235, 2024.
[16] *Fayzullo N., Sariyev S., Sherzodjon Y.* Analyzing the Effectiveness of Ansambl Methods in Solving Multi-Class Classification Problems. SmartIndustryCon 2025, Sochi, Russia, pp. 788–793. https://doi.org/10.1109/SmartIndustryCon65166.2025.10986248
[17] *Patil S.V., Wankhade N.R., Bagal S.B., Patel M.T.* Water Quality Analysis and Prediction Using Machine Learning (SVM included). Journal of Information Systems Engineering and Management.
[18] *Seviya T., Hakim L.* Water Quality Classification Using Support Vector Machine with PSO-Based Parameter Optimization. Information Technology Education Journal, 2025.
[19] *Uddin B., Marimin, et al.* Integration of RF-SVM Hybrid Model for Water Quality Prediction. Journal of Information Systems Engineering and Management, 2025.

Received: January 12,2026

Citation: *Nazarov F., Salimova M., Shukurov B.* 2026. Intelligent algorithms to assist in making water quality decisions based on artificial intelligence models. *International journal of theoretical and applied issues of digital technologies*. Volume 9, Issue 2, pp. 54-63. https://doi.org/10.62132/ijdt.v9i2.376.

ИНТЕЛЛЕКТУАЛЬНЫЕ АЛГОРИТМЫ ДЛЯ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ ПО КАЧЕСТВУ ВОДЫ НА ОСНОВЕ МОДЕЛЕЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

[+] *Назаров Ф.*[1], *Салимова М.*[1], *Шукуров Б.*[1]

[1] Самаркандский государственный университет имени Шарофа Рашидова, Самарканд, Узбекистан

Аннотация. В данной статье рассматриваются методы анализа состава воды и возможности использования современных технологий. В исследовании были сопоставлены традиционные физические, химические и биологические методы оценки качества воды, а также измерения, выполняемые с использованием сенсорных технологий. Была рассмотрена возможность быстрого и точного определения параметров воды с помощью таких устройств, как Exo2 sonde и TDS сенсор. Кроме того, было показано, что алгоритмы, основанные на искусственном интеллекте, могут классифицировать образцы воды при оценке и анализе качества воды, оценивать их по различным параметрам и группировать по категориям. Также был проведён анализ применения моделей Decision Tree, Random Forest и SVM. Результаты исследования рассмотрели возможности использования систем мониторинга воды как эффективного инструмента, повышающего эффективность и точность.

Ключевые слова: WQI, сенсорные технологии, искусственный интеллект, машинное обучение, decision tree, random forest, SVM, min-max scaling, алгоритм SMOTE.