BIO Web of Conferences **130**, 03016 (2024) https://doi.org/10.1051/bioconf/202413003016
BFT-2024

Machine learning in environmental monitoring: the case of water potability prediction

Vasiliy Orlov[1,2*], *Anatoliy* Kukartsev[2,3], *Ilya* Panfilov[2,3], *Anastasia* Kozlova[2,3], and *Andrey* Boyko[2,3]

[1] Institute of Business Process Management, Department of Digital Management Technologies, Siberian Federal University, Krasnoyarsk, Russia
[2]Bauman Moscow State Technical University, Artificial Intelligence Technology Scientific and Education Center, Moscow, Russia
[3]Reshetnev Siberian State of Science and Technology, Krasnoyarsk, Russia

Abstract. Ensuring the quality of drinking water is a critical public health issue, especially in the face of increasing industrial pollution and climate change. This study explores the application of machine learning techniques, specifically XGBoost, to predict water potability based on physical and chemical parameters. Initial experiments with a baseline XGBoost model demonstrated moderate success in classification, with particular difficulty in accurately identifying potable water, which is often underrepresented in the dataset.To address this challenge, the SMOTE technique was applied to balance the dataset, resulting in improved recall for the potable water class. Additionally, hyperparameter optimization via RandomizedSearchCV further enhanced the model's performance, albeit modestly. Reducing the model's complexity by selecting the most significant features led to a more efficient model with reduced computational overhead, while maintaining comparable accuracy.The findings of this study indicate that an optimized XGBoost model, supported by data preprocessing and feature selection, can be effectively integrated into automated water quality monitoring systems. Such a system would enable real-time detection of water quality deviations, facilitating prompt corrective actions to protect public health. This approach not only demonstrates the applicability of machine learning in environmental monitoring but also highlights the potential for broader application in resource management and public safety.

1 Introduction

Water is a vital resource for all living organisms, and its quality plays a crucial role in ensuring public health and well-being. With increasing population growth and industrial development, the quality of drinking water has become an increasingly pressing issue. Water contamination with chemical and biological substances can lead to serious health consequences, including acute and chronic diseases. Therefore, determining water potability is one of the key challenges in the fields of public health and environmental protection [1, 2].

* Corresponding author: vasi4244@gmail.com

In recent years, significant attention has been directed towards the development and application of various methods for water quality analysis, including traditional laboratory methods as well as modern approaches based on machine learning and data analysis. These methods allow for more efficient and accurate determination of the chemical composition of water and the prediction of its potability based on multiple parameters [3, 4].

The aim of this study is to develop and evaluate a machine learning model for predicting water potability based on chemical and physical parameters, such as pH, hardness, solids content, and others. The study explores various machine learning algorithms, including Random Forest and XGBoost, as well as data processing techniques such as class balancing using SMOTE and the selection of the most significant features [5, 6].

This work is aimed at improving existing methods for predicting water quality and can make a significant contribution to public health by providing more accurate and effective tools for monitoring water quality [7, 8 ,9].

2 Materials and methods

2.1 Data Sources

In this study, data on drinking water quality provided on the Kaggle platform were used. The dataset contains information on various physico-chemical parameters of water that can affect its potability. Specifically, the dataset includes the following key parameters [10, 11, 12]:

- pH: The level of water acidity, which can vary depending on natural and anthropogenic factors.
- Hardness: The overall concentration of calcium and magnesium in the water, which affects its taste and can influence health.
- Solids: The total amount of dissolved inorganic substances in the water, which can vary depending on the water source.
- Chloramines: The concentration of chloramines, used as a disinfectant in tap water.
- Sulfate: The concentration of sulfates, which may be present in water as a result of natural processes or contamination.
- Conductivity: The ability of water to conduct electric current, which depends on the concentration of dissolved salts.
- Organic Carbon: The amount of organic carbon, which can be an indicator of organic pollution.
- Trihalomethanes: The total amount of trihalomethanes, which can form as a byproduct of water chlorination.
- Turbidity: A measure of water turbidity, which can indicate the presence of suspended particles and microorganisms.
- Potability: The target variable indicating whether the water is potable (1) or not potable (0).

2.2 Data Description

The dataset used in this study contains information on various physical and chemical parameters of water that can affect its potability. The data analyzed consists of 3,276 records, each corresponding to a single water sample collected from different sources [13].

Table 1. Descriptive statistics

Parameter	Description	Data Type	Value Range
pH	The level of water acidity. The optimal range for drinking water is 6.5-8.5.	Float	0.0 - 14.0
Hardness	The hardness of water, measured in mg/L, representing the concentration of calcium and magnesium.	Float	47.43 - 323.12
Solids	The total concentration of dissolved solids in water, measured in ppm.	Float	320.94 - 61227.20
Chloramines	The concentration of chloramines used for water disinfection, measured in ppm.	Float	0.35 - 13.13
Sulfate	The concentration of sulfates in water, measured in mg/L.	Float	129.0 - 481.0
Conductivity	The conductivity of water, measured in microsiemens per centimeter (μS/cm).	Float	181.48 - 753.34
Organic Carbon	The amount of organic carbon in water, measured in mg/L.	Float	2.2 - 28.3
Trihalomethanes	The concentration of trihalomethanes, which can form during chlorination, measured in μg/L.	Float	0.74 - 124.0
Turbidity	A measure of water turbidity, measured in NTU units.	Float	1.45 - 6.74
Potability	The target variable indicating water potability (1 — potable, 0 — non-potable).	Integer	0 or 1

2.3 Model Training and Evaluation

Before training the models, the data was split into training and test sets in an 80:20 ratio, which is a standard practice for evaluating model performance. To minimize the impact of differences in feature scales, the data was normalized, which improved the convergence and accuracy of the models [14, 15].

Additionally, to address the issue of imbalanced data, where the number of samples in one class significantly exceeds the number of samples in the other, the SMOTE (Synthetic Minority Over-sampling Technique) method was applied. This method artificially increases the number of minority class examples by creating synthetic samples, which in turn enhances the model's ability to recognize both classes [16, 17, 18].

3 Results

3.1 Correlation Analysis

A correlation analysis was conducted to assess the relationships between various physico-chemical parameters of water and their impact on water potability. The correlation matrix (Figure 1) provides a quantitative measure of the linear relationships between pairs of features, allowing for the evaluation of their potential influence on predictive models [20, 21, 22].

 https://doi.org/10.1051/bioconf/202413003016

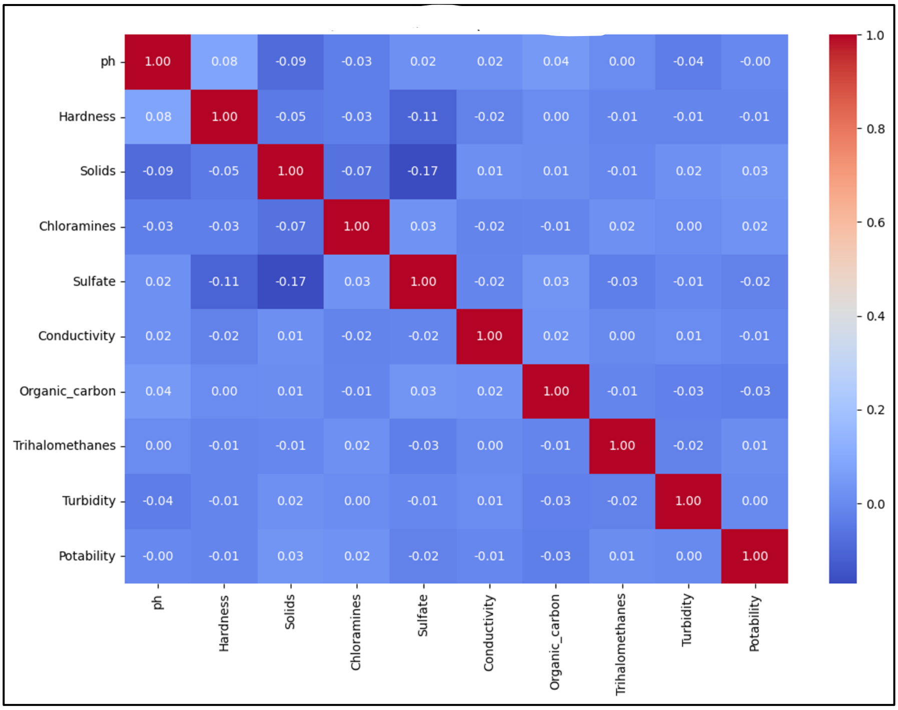

Fig. 1. Correlation matrix of water quality parameters

Key Findings:

- Weak relationships between features: The analysis showed that most features exhibit weak correlations with each other, indicating relative independence among the parameters. The maximum correlation values between features do not exceed |0.17|, suggesting weak linear relationships. For example, the highest negative correlation was found between Sulfate and Solids (-0.17), while a positive correlation was observed between Hardness and pH (0.08).
- Weak correlation with the target variable: The target variable, Potability, shows very low correlation with the other parameters. The highest correlation with the target variable was observed with the Solids parameter (0.03), which also indicates a weak relationship. This suggests that none of the individual features have a significant linear impact on water potability.
- Implications for model building: The lack of strong correlations between features and the target variable highlights the complexity of the classification task and the need for advanced machine learning models. Linear models, such as logistic regression, may prove insufficient, supporting the use of methods capable of capturing nonlinear interactions between features, such as gradient boosting (e.g., XGBoost).

The results of the correlation analysis indicate that models capable of effectively handling data with weak linear relationships between features are essential for predicting water potability. These findings motivated the use of models that can identify complex dependencies and interactions between features, which was considered in the subsequent selection and tuning of models in this study.

3.2 Experiment with the Baseline XGBoost Model

The objective of this experiment was to assess the performance of a baseline XGBoost model for predicting water potability without the use of additional data processing methods, such as class balancing, feature selection, or hyperparameter optimization. This experiment serves as a starting point for further model improvements [23, 24].

The results of the experiment are presented through a classification report and a confusion matrix, which provide deeper insights into how the model handles the classification task.

The XGBoost model was initialized with default hyperparameters, including 100 decision trees (n_estimators), a maximum tree depth of 3 (max_depth), and a learning rate of 0.1 (learning_rate).

As shown in Table 2, the baseline XGBoost model achieved a precision of 0.70 for class 0 and 0.55 for class 1. The recall was 0.79 and 0.42, respectively, indicating that the model is better at recognizing non-potable water (class 0) than potable water (class 1).

Table 2. Classification Report for the Baseline XGBoost Model

Class	Precision	Recall	F1-Score	Support
0	0.70	0.79	0.74	412
1	0.55	0.42	0.48	244

The confusion matrix, presented in Figure 2, shows that the model correctly classified 327 samples of class 0 and 103 samples of class 1. However, the model mistakenly classified 85 samples of class 0 as class 1 and 141 samples of class 1 as class 0. This confirms that the model is better at recognizing non-potable water (class 0) than potable water (class 1).

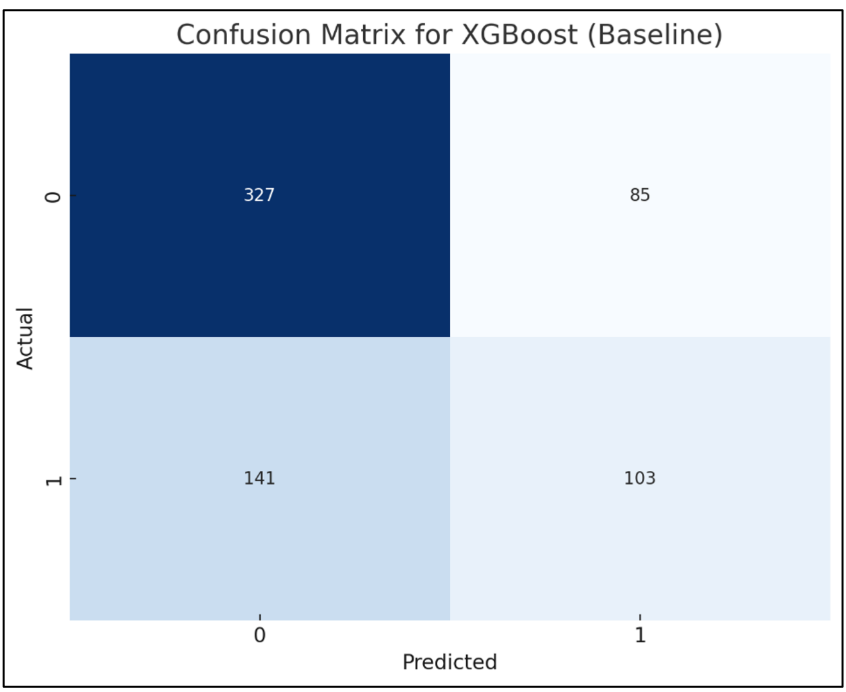

Fig. 2. Confusion Matrix for XGBoost (Baseline)

The baseline XGBoost model showed limited results, with an accuracy of 66% and a ROC-AUC score of 0.69. While the model performs better at recognizing non-potable water, it struggles with correctly classifying potable water. These results highlight the need for further improvements, such as class balancing, hyperparameter optimization, and feature selection, to enhance overall accuracy and the model's ability to distinguish between classes.

3.3 Application of XGBoost Using SMOTE

The objective of this experiment was to improve the performance of the XGBoost model by balancing the classes in the training set using the SMOTE (Synthetic Minority Over-sampling Technique) method. This approach was implemented to enhance the model's ability to recognize the minority class (potable water), which was underrepresented in the data and exhibited low recall in the baseline model [25].

To address the class imbalance present in the original data, the SMOTE method was applied. This technique generates synthetic samples for the minority class (potable water), allowing the model to better learn from these examples and improve recall for class 1.

After balancing the data, the XGBoost model was trained on the new training set using the same baseline hyperparameters as in the first experiment :

- n_estimators: 100;
- max_depth: 3;
- learning_rate: 0.1.

The results of the experiment using SMOTE are presented through a classification report and a confusion matrix, allowing for an analysis of the impact of class balancing on the model's performance.

As shown in Table 3, the application of SMOTE led to an improvement in recall for class 1 (potable water) from 0.42 to 0.57, indicating a more balanced performance by the model.

Precision for class 1 remained at 0.54, while precision for class 0 increased to 0.74. The F1-score for both classes also improved, reflecting an overall enhancement in the model's performance.

Table 3. Classification Report for XGBoost Model Using SMOTE

Class	Precision	Recall	F1-Score	Support
0	0.74	0.72	0.73	412
1	0.54	0.57	0.55	244

The confusion matrix, presented in Figure 3, shows that the model correctly classified 295 samples of class 0 and 138 samples of class 1. The application of SMOTE reduced the number of misclassifications for class 1, indicating an improvement in the recognition of potable water.

https://doi.org/10.1051/bioconf/202413003016

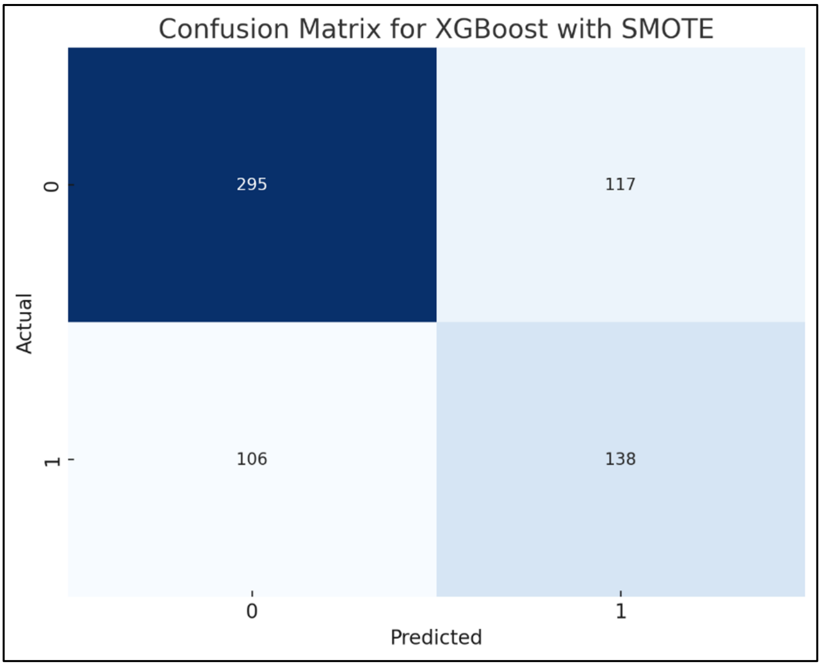

Fig. 3. Confusion Matrix For XGBoost With SMOTE

The application of SMOTE for class balancing improved recall for class 1 (potable water), demonstrating the model's enhanced ability to recognize this class. Although the overall accuracy of the model remained at 66%, the improvement in both F1-score and recall for class 1 makes this model more balanced compared to the baseline version. This confirms the effectiveness of using data balancing techniques when dealing with imbalanced datasets.

3.4 Hyperparameter Optimization of the XGBoost Model

The objective of this experiment was to enhance the performance of the XGBoost model by optimizing its hyperparameters. Hyperparameter optimization allows for fine-tuning the model to achieve the best possible results based on the available data. In this case, the RandomizedSearchCV method was employed to find the optimal hyperparameter values [26].

The RandomizedSearchCV method was used for hyperparameter optimization, which explores the hyperparameter space based on predefined ranges and random combinations. This approach allows for an efficient search for the best configuration that can potentially improve the model's accuracy and overall performance.

The search space included the following hyperparameters:

- n_estimators: 100, 200, 300, 400, 500;
- max_depth: 3, 4, 5, 6, 7, 8, 10;
- learning_rate: 0.01, 0.05, 0.1, 0.2;
- subsample: 0.6, 0.7, 0.8, 0.9, 1.0;
- colsample_bytree: 0.6, 0.7, 0.8, 0.9, 1.0;
- gamma: 0, 0.1, 0.2, 0.3, 0.4;
- reg_alpha: 0, 0.1, 0.5, 1;
- reg_lambda: 0, 0.1, 0.5, 1.

 https://doi.org/10.1051/bioconf/202413003016

The RandomizedSearchCV method was configured to perform 50 iterations with 3-fold cross-validation to identify the optimal combination of hyperparameters.

The best hyperparameters found as a result of the search were as follows:

- subsample: 0.9;
- reg_lambda: 0.5;
- reg_alpha: 0.1;
- n_estimators: 200;
- max_depth: 10;
- learning_rate: 0.05;
- gamma: 0.3;
- colsample_bytree: 1.0.

These parameters were used to train the model on the same data as in the previous experiments.

The results of the experiment are presented through a classification report and a confusion matrix, allowing an assessment of the impact of hyperparameter optimization on the model's performance.

As shown in Table 4, the optimized XGBoost model demonstrated a slight improvement in performance compared to the baseline model. The model achieved a precision of 0.72 for class 0 and 0.54 for class 1. Recall was 0.74 and 0.51, respectively. The F1-score remained at 0.73 for class 0 and 0.53 for class 1.

Table 4. Classification Report for the Optimized XGBoost Model

Class	Precision	Recall	F1-Score	Support
0	0.72	0.74	0.73	412
1	0.54	0.51	0.53	244

The confusion matrix, presented in Figure 4, shows that the model correctly classified 305 samples of class 0 and 125 samples of class 1. Hyperparameter optimization slightly improved the model's ability to correctly classify both classes.

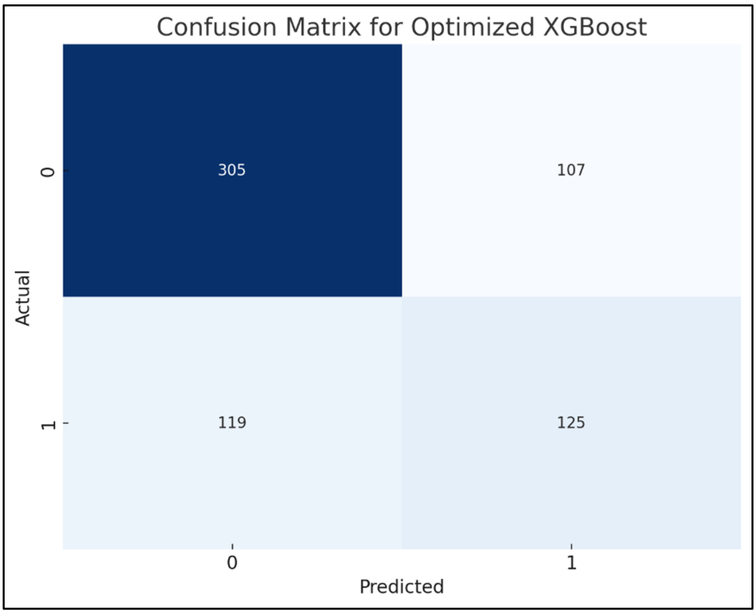

Fig. 4. Confusion Matrix for optimized XGBoost

https://doi.org/10.1051/bioconf/202413003016

The hyperparameter optimization of the XGBoost model led to a slight improvement in precision and recall, particularly for class 1 (potable water). However, the overall results indicate that despite these improvements, the model still struggles to accurately identify potable water, suggesting the need for further enhancements, such as employing more complex models or combining multiple methods. Nevertheless, this experiment demonstrated that hyperparameter optimization can play a significant role in enhancing model performance.

3.5 Application of XGBoost Using the Most Significant Features

The objective of this experiment was to investigate the impact of selecting the most significant features on the performance of the XGBoost model. This approach aimed to reduce the complexity of the model and enhance its interpretability, while also examining whether using a smaller number of features would lead to an improvement or deterioration in prediction accuracy [27].

Based on the analysis of feature importance obtained from the XGBoost model, the 5 most significant features were selected for further model training (Figure 5).

The feature selection was conducted with the goal of utilizing only those variables that contribute the most to the model's predictions. This is expected to reduce the model's complexity and improve its computational efficiency, potentially leading to faster training and prediction times. Additionally, by focusing on the most impactful features, the model's interpretability is enhanced, making it easier to understand the relationships between the selected features and the target variable.

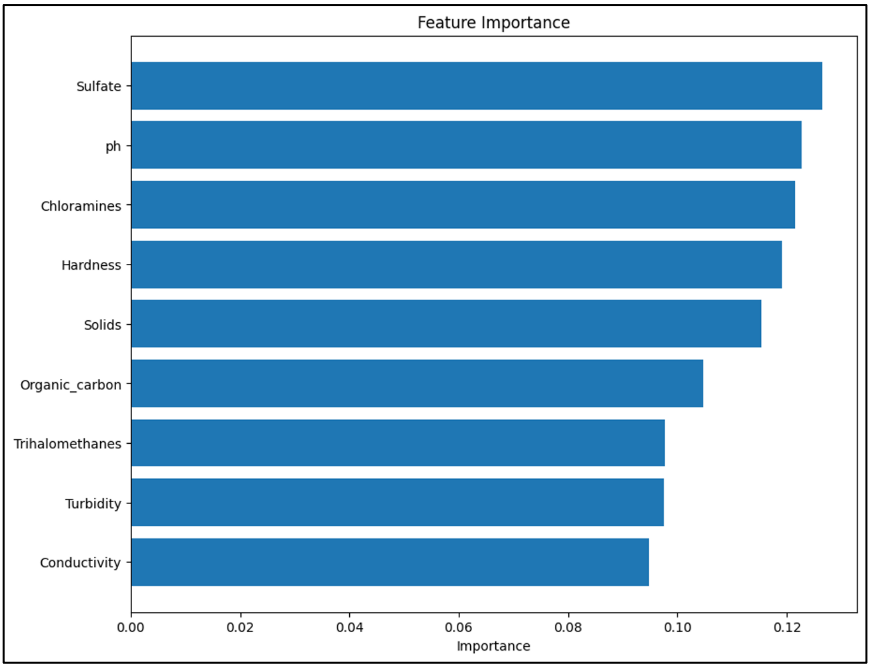

Fig. 5. Feature importance

The same optimized hyperparameters that were identified in the previous experiment were used for training the model:

- subsample: 0.9;
- reg_lambda: 0.5;

- reg_alpha: 0.1;
- n_estimators: 200;
- max_depth: 10;
- learning_rate: 0.05;
- gamma: 0.3;
- colsample_bytree: 1.0.

The results of the experiment are presented through a classification report. As shown in Table 5, the XGBoost model using the top features achieved a precision of 0.70 for class 0 and 0.55 for class 1. The recall was 0.79 and 0.43, respectively. These results indicate that reducing the number of features did not lead to a significant improvement in the model's performance. However, it is important to note that using fewer features reduced the model's complexity, which could be beneficial for computational efficiency and interpretability.

Table 5. Classification Report for XGBoost Model with Top Features

Class	Precision	Recall	F1-Score	Support
0	0.70	0.79	0.74	412
1	0.55	0.43	0.48	244

The confusion matrix, presented in Figure 6, shows that the model correctly classified 326 samples of class 0 and 104 samples of class 1. This confirms that feature selection did not lead to a significant increase in prediction accuracy; however, it did reduce the model's complexity without substantial losses in performance.

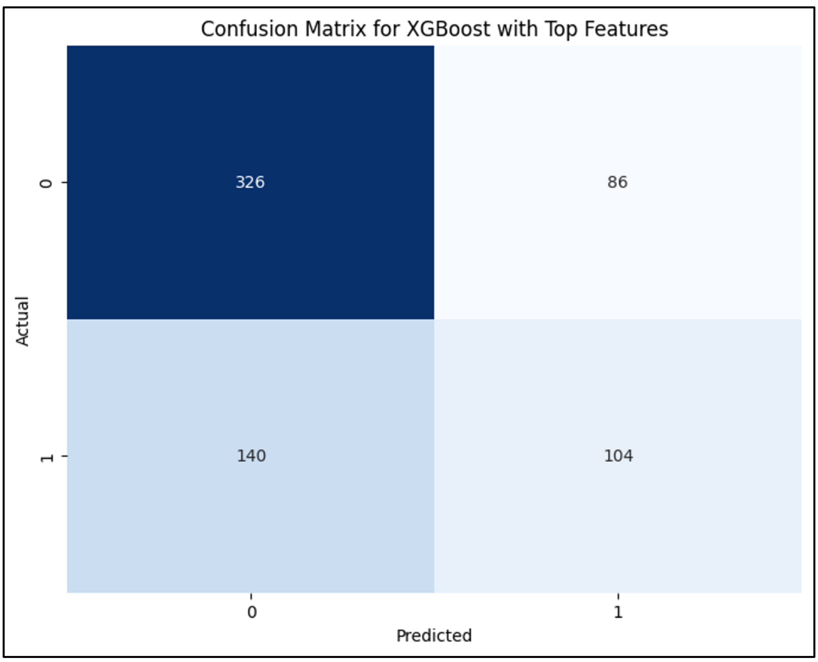

Fig. 6. Confusion Matrix for XGBoost Model with Top Features.

The results of the experiment demonstrated that using only the most significant features allowed for a reduction in model complexity without significant losses in performance. While the model's accuracy and ability to distinguish between classes remained consistent with previous experiments, the advantages in speed, interpretability, and resistance to overfitting make this approach valuable in applications where these aspects are crucial. This confirms the appropriateness of feature selection in scenarios where a balance between performance and efficiency is required.

4 Discussion

Throughout this study, several experiments were conducted using various approaches to develop and enhance the XGBoost model for predicting water potability. This section discusses the key findings from these experiments, compares the different approaches, analyzes their advantages and disadvantages, and suggests potential directions for future research.

4.1 Baseline XGBoost Model

The initial experiment with the baseline XGBoost model demonstrated that, despite the model's ability to handle complex and interrelated data, it showed limited results. The model achieved an accuracy of 66% and a ROC-AUC score of 0.69, indicating moderate capability in distinguishing between potable and non-potable water. The low recall for class 1 (potable water) highlights the model's difficulty in recognizing the minority class, a common issue with imbalanced datasets.

4.2 Application of SMOTE for Class Balancing

The use of the SMOTE method for class balancing improved the recall for class 1, increasing it from 0.42 to 0.57. This indicates that the model became better at recognizing potable water. However, the overall accuracy of the model remained at 66%, and the ROC-AUC score did not show significant changes. Thus, while SMOTE enhanced the model's ability to classify the minority class, it did not lead to substantial improvements in other metrics.

4.3 Hyperparameter Optimization

Hyperparameter optimization of the XGBoost model using RandomizedSearchCV resulted in slight improvements in model performance. Optimal parameters, such as reduced learning rate and increased tree depth, improved accuracy to 72% for class 0 and 54% for class 1. Despite this, the overall results indicate that the model still performs better at recognizing non-potable water than potable water. This underscores the importance of further tuning and potentially applying other methods to enhance the model's effectiveness.

4.4 Reducing Model Complexity with Top Features

Using only the most significant features allowed for a reduction in model complexity without significant losses in performance. While accuracy and other metrics remained at similar levels to previous experiments, reducing the model's complexity has its advantages. A simpler model trains faster, makes predictions quicker, and is less prone to overfitting. This approach is particularly appealing in scenarios where speed and interpretability are critical, such as in real-time systems or environments with limited computational resources.

5 Conclusion

Ensuring the quality of drinking water is one of the key challenges in public health and environmental protection. With increasing industrial impact and climate change, the need for reliable and effective water monitoring methods is becoming more urgent. In this study, we demonstrated how machine learning methods, particularly XGBoost, can be used for the automated prediction of water potability based on its physical and chemical parameters [28].

The experiments revealed that while the baseline XGBoost model is effective in classifying water by its potability, it struggles with recognizing the potable water class, which often constitutes the minority of the dataset. To address this issue, the SMOTE method was applied, which improved class balancing and increased the accuracy of potable water classification. This is particularly important in practical applications where errors in recognizing potable water can have serious consequences [29].

Additionally, hyperparameter optimization showed that fine-tuning the model can slightly improve its performance, which is critical for real-world monitoring systems that require high accuracy. Reducing model complexity by selecting the most significant features led to decreased computational costs without significant losses in accuracy, making the model more suitable for real-time use [30].

Thus, the study's results demonstrate that the XGBoost model, combined with data preprocessing and optimization methods, can be effectively applied for automated water quality monitoring. Such a system can be integrated into existing water supply infrastructures, ensuring timely detection of water quality deviations and allowing for rapid corrective actions to protect public health. In the future, this model can also be adapted for other environmental monitoring tasks, making it a versatile tool for natural resource management [31].

References

1. Bosikov I. I. et al. Modeling and complex analysis of the topology parameters of ventilation networks when ensuring fire safety while developing coal and gas deposits //Fire. – 2023. – T. 6. – №. 3. – C. 95.
2. Malozyomov B. V. et al. Study of Supercapacitors Built in the Start-Up System of the Main Diesel Locomotive //Energies. – 2023. – T. 16. – №. 9. – C. 3909.
3. Strateichuk D. M. et al. Morphological Features of Polycrystalline CdS1− xSex Films Obtained by Screen-Printing Method //Crystals. – 2023. – T. 13. – №. 5. – C. 825.
4. Martyushev N. V. et al. Review of Methods for Improving the Energy Efficiency of Electrified Ground Transport by Optimizing Battery Consumption //Energies. – 2023. – T. 16. – №. 2. – C. 729.
5. Shutaleva A. et al. Sustainability of Inclusive Education in Schools and Higher Education: Teachers and Students with Special Educational Needs //Sustainability. – 2023. – T. 15. – №. 4. – C. 3011.
6. Kukartsev V. A. et al. Study of the Influence of the Thermal Capacity of the Lining of Acid Melting Furnaces on Their Efficiency //Metals. – 2023. – T. 13. – №. 2. – C. 337.
7. Martyushev N. V. et al. Provision of Rational Parameters for the Turning Mode of Small-Sized Parts Made of the 29 NK Alloy and Beryllium Bronze for Subsequent Thermal Pulse Deburring //Materials. – 2023. – T. 16. – №. 9. – C. 3490.
8. Rezanov V. A. et al. Study of melting methods by electric resistance welding of rails //Metals. – 2022. – T. 12. – №. 12. – C. 2135.
9. Bashmur K. A. et al. Biofuel technologies and petroleum industry: synergy of sustainable development for the Eastern Siberian Arctic //Sustainability. – 2022. – T. 14. – №. 20. – C. 13083.

10. Kolenchukov O. A. et al. Experimental study of oil non-condensable gas pyrolysis in a stirred-tank reactor for catalysis of hydrogen and hydrogen-containing mixtures production //Energies. – 2022. – T. 15. – №. 22. – C. 8346.
11. Tynchenko Ya. A. et al. Probabilistic analysis of pump reliability indicators using a neural network. MIAB. Mining Inf. Anal. Bull. 2024;(7-1):126-136. [In Russ]. DOI: 10.25018/0236_1493_2024_71_0_126.
12. Tynchenko Ya.A., Kukartsev V.V., Gladkov A.A., Panfilova T.A. Assessment of technical water quality in mining based on machine learning methods. Sustainable Development of Mountain Territories, 2024, vol. 16, no. 1, pp. 56–69. (In Russ.). DOI: 10.21177/1998-4502-2024-16-1-56-69
13. Kukartsev V. V., Kravtsov K. I., Tynchenko Ya. A., Panfilova T. A. Influence of mountain factors on salt excess and soil toxicity in mountain conditions. Sustainable Development of Mountain Territories. 2023;15(3):784-797
14. Yelemessov K.K. et al. Mathematical and computer modeling of gantry crane load-beam system oscillation / Sustainable Development of Mountain Territories. 2023;15(2): 450-461.
15. – DOI 10.21177/1998-4502-2023-15-2-450-461.
16. Brigida V. et al. (2024). Technogenic Reservoirs Resources of Mine Methane When Implementing the Circular Waste Management Concept. Resources, 13(2), 33.
17. Malozyomov B. V. et al. Determination of the Performance Characteristics of a Traction Battery in an Electric Vehicle. World Electric Vehicle Journal, 15(2), 64.
18. Sokolov A. A. et al. Ensuring uninterrupted power supply to mining enterprises by developing virtual models of different operation modes of transformer substations. MIAB. Mining Inf. Anal. Bull. 2023;(11-1): 278-291.
19. Gutarevich V. O. et al. Reducing oscillations in suspension of mine monorail track. MIAB. Mining Inf. Anal. Bull. 2023;(11-1):72-87.
20. Degtyareva K. et al. Finding patterns in employee attrition rates using self-organizing Kohonen maps and decision trees. In 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES) (pp. 1-6). IEEE.
21. Gladkov A. et al. Development of Requirements for AIS Aimed at Controlling High Turnover. In 2023 IEEE International Conference on Computing (ICOCO) (pp. 232-236). IEEE.
22. Zhilkina Y. et al. Strategy of introduction of information system in trade and logistics company. In E3S Web of Conferences (Vol. 458, p. 05016). EDP Sciences, 2023
23. Kukartsev V. V. et al. Advancements in network-based management systems for enhanced business services //E3S Web of Conferences. – EDP Sciences, 2023. – T. 460. – C. 07003.
24. Kozlova A. et al. Finding dependencies in the corporate environment using data mining. In E3S Web of Conferences (Vol. 431, p. 05032). EDP Sciences, 2023
25. Kukartsev V. et al. Control system for personnel, fuel and boilers in the boiler house. In E3S Web of Conferences (Vol. 458, p. 01010). EDP Sciences, 2023
26. Vasileva V. et al. Integration of automated information systems and architectural solutions in industrial enterprises. In E3S Web of Conferences (Vol. 458, p. 09021). EDP Sciences, 2023

27. Gladkov A. et al. Development of an automation system for personnel monitoring and control of ordered products //E3S Web of Conferences 458, 01007 (2023). – EDP Sciences, 2023.
28. Orlov, V. et al. Designing an information system to automate service management at the enterprise. In E3S Web of Conferences (Vol. 458, p. 09019). EDP Sciences.
29. Kolenchukov O.A. et al. Forecasting the technical condition of thermochemical reactor systems / SOCAR Proceedings No.1 (2023) 123-130
30. http://dx.doi.org/10.5510/OGP20230100814
31. Malozyomov B. V. et al. Substantiation of Drilling Parameters for Undermined Drainage Boreholes for Increasing Methane Production from Unconventional Coal-Gas Collectors //Energies. – 2023. – T. 16. – №. 11. – C. 4276.