

WaterCare: A Monitoring and Assessment System for Agricultural and Domestic Water Quality

¹C Giridhar, ²Angoth Lakshman, ³E Nikhil Reddy, ⁴P Srikanth

¹Student, ²Assistant Professor, ³Student, ⁴Student

¹Computer Science Engineering (Data Science),

¹CMR Technical Campus, Hyderabad, India

Abstract : The need to ensure that the water people consume and also use in agriculture is important but may take long. The general method of testing water involves laboratory testing which is not always as fast as quick decisions. This paper presents a machine-learned system called WaterCare, which analyzes the data on water quality and provides information on whether a sample can be utilized to drink or water the land. The system deals with water parameters which are pH, turbidity, hardness, conductivity and dissolved substances. These values are initially washed and sorted out and then they are subject to be trained on various models, such as, Logistic Regression, Random Forest and Support Vector machine. We compare the models so as to select the one providing the best results. The system displays the contribution of each parameter to the final decision other than the predictions, thus, is less cumbersome to comprehend. An easy to use web interface would allow the entry of water parameters and immediate predictions with a detailed visual output. Using data analysis rather than manual testing to analyze water quality, WaterCare provides a convenient and quicker method of testing the water quality. Such system may be helpful in the case of students, researchers, farmers, and so on.

Every person that requires a simple and realistic means of determining whether water is appropriate.

Keywords: Water Quality, Machine Learning, Drinking water, irrigation, contribution of features, prediction, data analysis.

I.

INTRODUCTION

Supplied clean water is a fundamental requirement to our life, agriculture, and maintaining the environment [12]. Factories, increasing cities, and increased pollution have transformed most water sources over time [1, 5]. Due to these changes, it is difficult to state whether water is safe to consume or it is good to grow crops. Lab tests are the standard method of assessing the quality of water [10]. They are accurate tests, however they are expensive, time consuming and require a great deal of effort [9, 10]. That makes them not so helpful when they are used commonly or when they are to cover long distance. This is why we require computer techniques by which we can analyze water information quick and fast and assist us in making swift choices [1, 3].

At this time, data analysis tools and machine learning have become better [1, 3]. They allow us to construct work systems that are automatically interacting with environmental data [2]. These models are able to identify trends, visualize correlation among various measurements of water and cluster samples accurately [3, 6]. This is beneficial when we have too many water properties such as pH, hardness, turbidity, conductivity, and dissolved substances where one is interested in knowing whether the water could be consumed to drink or used to farm [1, 6]. The data-driven classification of water provides repeatable stable outcomes, hence making us less dependent on rigorous lab experiments and achieve greater efficiency [1, 2].

WaterCare is constructed on the basis of data as a machine learning-solving data tool that uses data to determine whether water can be consumed or used to farm. It begins with cleaning and sorting of the water data. It will then train numerous prediction models and select the most appropriate. It also demonstrates which measurements contribute to the final decision the most, thus, users are able to learn how the system functions. The interface is user-friendly and easy as users can view the results on it, read explanations, and post questions in real-time. This renders the entire assessment comprehensible and user friendly.

WaterCare combines good data preparation, comparing models, and explaining it, creating an efficient, more economical method of determining the quality of water [1, 3]. It may assist in better surveillance of the environment and the resource usage

II. LITERATURE SURVEY

2.1 Advanced Monitoring Technologies and IoT Integration (2019-2021)

Water quality monitoring is the one of the most important environmental concern around the globe. Quality of water is represented in terms of physical, chemical and biological characteristics of which ions and chemical compounds are the most important. Scholars all over the world are integrating Internet of Things with augmented technologies in the computational field to enhance the monitoring capabilities of water quality. Foschi et al. [2] enhanced the soft sensor for predicting Escherichia coli concentration from traditional monitoring parameters. Moreover, Akhtar et al. [3] suggested characterization on the traditional WQI processes by integrating the multi criteria decision making (MCDM) based processes which allow for a simple procedure of calculation with accurate assessment results. Similarly, Uddin et al. [1] reviewed the various WQIs and their accuracy in evaluating surface water quality under different situations and different locations. Following are the various technologies integrated for.

2.2 Water Quality Index Development and Sustainable Development Goals (2020-2021)

In recent years, diverse water quality indices and methods have been advanced by different researchers. Uddin et al. [1] discussed different models indicating water quality indices used to assess the quality of surface water in different environments and geographical regions. According to the paper presented by Ho and others [5] a case study carried out in Belgium illustrates how water research can support sustainable development practice. The suggested approach that uses water as the focal point for case study development is essential to use water research to support sustainable development practice. Banda and Kumarasamy [6] create a universal water quality index (UWQI) for river catchments in South Africa. Ortiz et al. analyzed experimental vs simulation communication technologies of vehicular and environmental monitoring applications.

2.3 IoT and Communication Technologies for Environmental Monitoring (2017-2019)

The advancement of communications and computing technologies has created ample opportunity to design scalable water quality monitoring networks. Ma et al., [7], designed an IoT-based fog computing model for real-time data processing and analysis. Raza et al., [8], provides an overview of low power wide area networks (LPWAN) for environmental monitoring systems. These technologies can allay concerns about the scalability and cost of environmental monitoring.

2.4 Comprehensive Water Quality Monitoring Network Design (2008-2016)

It's crucial to have a strategic design of the network and implementation of international standards for the success of water quality monitoring. Chilundo et al. [9] built a water quality observation network in Mozambique's Limpopo river basin. A more acceptable framework for positioning the monitoring station in the river basin has been developed. Monitoring stations are placed to maximize data coverage and minimize costs of operation. Lee et al [11] explained the significance of water quality monitoring for the achievement of sustainable development goals. Monitoring of Water Quality Data-Driven has become need for WRM. To adequately plan and develop effective monitoring of such huge water systems, it is essential to very carefully consider the geographical, hydrological and resource aspect.

III. PROPOSED METHODOLOGY

3.1 Dataset Acquisition and Pre-Processing:

WaterCare gathers water-quality information on both the sources of water and documents maintained by the institutions. The measurements of the physical and chemical required are present in the data. Once the data has been collected, a individual dataset is thoroughly analyzed and cleaned to make it uniform, accurate and fit to be analyzed. The cleaning process imparts lost data, mitigates the impact of outliers, normalizes values whenever necessary, converts category into numbers and organizes all the fields in a lucid format that classification models find it easy to utilize.

Additional measures are made to enhance reliability of data by eliminating anomalous data through techniques of noise-reduction. Statistical summaries indicate the distribution of the important parameters, such as pH, hardness, solids, conductivity, and turbidity, among the data. Standardizations of the data to expose valuable patterns are performed by applying feature transformations such as scaling, checking of parameters correlations and elimination of the less useful dimensions. Once all these have been done, the resulting clean dataset gets to be the primary data to be used in training models, where performance is to be evaluated and results shall be compared throughout the entire classification procedure.

3.2 Model Calibration and Class-Specific Classification Strategy

Upon the data preparation, WaterCare develops models which determine whether one should consume the water sample or it should be tapped into irrigation. The system errs on more than one algorithm for use including, but not limited to, Logistic Regression, Random Forest, Support Vector Machine, and some simple classifiers. The multiplicity of methods allows observing how both models process the same data and simplifies the process of locating the model that works the best.

At this step, the models are adjusted very closely on the data. The data is divided into training and testing sample, the hyperparameters are varied by altering them and the models are tested repeatedly with various performance measures. The complexity is determined by factors such as the size of the dataset, the extent to which the parameter values differ as well as the distribution of the classes. This allows both drinking water evaluation and irrigation suitability to have equal predictions by each algorithm adapting itself naturally to the data.

Accuracy, precision, recall, and F1 -score are the standard metrics that are calculated to see a fair and useful evaluation per model. This facilitates transparency in comparison. Adapting the models to the specifications of the data, WaterCare develops the stable and reliable classification process that facilitates continuous improvement and the long-term reliability in the analytical level.

3.3 Interactive Prediction Workflow and Real-Time Analytical Feedback:

The interface of the system is simple as people can enter the water test values and request a check. Simply drop an item onto the models and then as soon as you do an outcome will appear in front of your eyes. Tools, such as examining the significance of each parameter, quantifying the influence of every parameter, and perusing the confidence of the forecast, contribute to the real-time solution.

Upon the result coming people receive a clear feedback explanation of the outcome, any odd reading on the reading is flagged and the primary reasons of the decision are displayed. There could also be some notes that are helpful giving details of how the various parameters fit together during the evaluation.

It is also possible to get such detailed reports through the interface that indicate accuracy trends, model behavior, and performance outcomes based on what was previously requested by the user. By this interactive workflow, the users are taught more on water quality assessment and are able to make environmental decisions.

3.4 Performance Evaluation Engine and Parameter Influence Analysis:

Once a classification cycle has completed, the performance engine looks at the performance of the models. It computes statistics that depict their behavior including general accuracy, number of errors, confusion, and a side by side comparison of various algorithms. Through this review, it is possible to discover which models are being consistent and which require further improvements.

In addition to measuring the performance, the engine also includes a feature-importance analysis based on such techniques as recursive elimination, permutation testing, and impurity ranking. Such techniques demonstrate the influence of the parameters on the predictions the most and the contribution of each attribute. Observing these patterns, experts get to know how various properties influence findings and what factors invariably affect the model.

Based on its learning, the engine can make recommendations on how to be improved, e.g. by fine-tuning its models, by changing parameter settings, or by increasing the scale of its validation tests in case of weaknesses in its data. Such recommendations will assist in maintaining the process positive and ensure that the system remains capable and flexible to future information or other parametric requirements.

3.5 System Architecture:

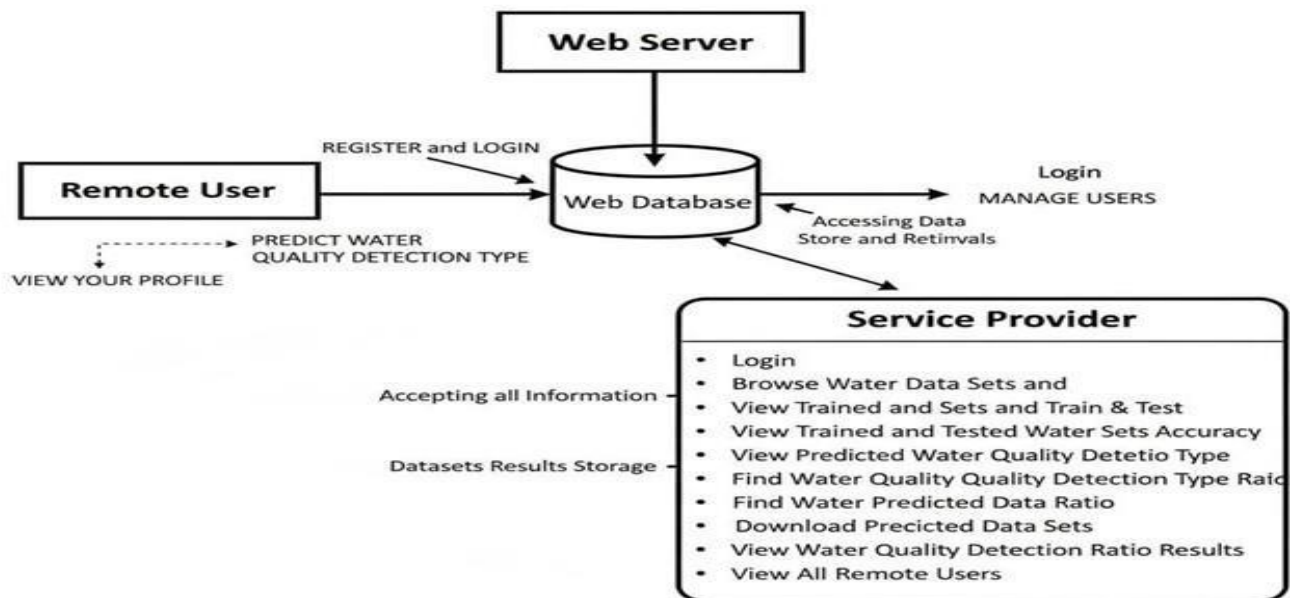


fig. 1: system architecture

WaterCare framework has a clientserver model as shown in the above fig. 1 in which all uses, data analysis and data interactions take place under a single central web venous systems. The system will be constructed using three primary components of it such as Remote User, Web Server and Service Provider that have a common database where data is stored, retrieved and processed with ease.

Remote User interface also allows individuals to create accounts, register/log in, and input water-quality values to receive predictions. Users have the facility to request classification, view previous results and update their profiles without any hard work. All requests begin at the user level, which is sent to the web server, validated at the web server and channeled to the appropriate analysis module.

The primary aspect of the system is the Web Server. It processes all the incoming user requests, mediates between the components and ensures that all the dataset activities are done. It is connected with the Web Database where clean water datasets, models outputs model, user information, prediction history, and accuracy will be saved. Upon the next query, a user is provided with a simple answer after the trained models, prediction steps, and the server perform the work. It also takes administrative operations provided by the Service Provider, including issuing permission to users, updating datasets, and performance monitoring of the system as a whole.

The administrative and analytic module is the Service Provider module. The provider can through a dashboard upload and manage datasets, train and test models, evaluate performance, calculate accuracy and view classification results of various algorithms. The provider is also able to see the detect rates, download prediction outcomes and to see the Remote User activity. This element maintains dataments up to date, models operating correctly and the system managed.

The system has its central storage in Web Database. It contains training and test data sets, prediction results record, user passwords and accuracy statistics. The database will be designed in an efficient manner of storage and easy retrieval of the data, thus making nothing to be lost and nothing to run awry.

IV. RESULTS AND DISCUSSION

4.1 Results of WAMP server

The application was run on WAMP server and all needed services were started without any problem. No errors were present in the beginning of the system. When the server was online, we were able to access the WaterCare platform on local host. Navigation was straightforward and one could confirm that everything worked. The information about the project was entered into the application without issues which meant that the system was capable of reading the chemical and physical values and preparing them to be analyzed. The database connections and the files were also located in the system which indicated that the server operation and file management was configured correctly as shown in **fig. 2**. The information that we experienced on the running made us confident that the hosting environment and data integration were performing well. It demonstrated that the platform was prepared to do additional processing and analysis without any technical issues.

```
Microsoft Windows [Version 10.0.26200.6901]
(c) Microsoft Corporation. All rights reserved.

C:\Users\girid\OneDrive\Desktop\major\WaterNet\WaterNet\waternet>python manage.py runserver
Performing system checks...

System check identified no issues (0 silenced).

Not checking migrations as it is not possible to access/create the django_migrations table.
October 29, 2025 - 14:27:34
Django version 1.11.6, using settings 'waternet.settings'
Starting development server at http://127.0.0.1:8000/
Quit the server with CTRL-BREAK.
```

fig. 2: Running the server

4.2 Results of Graphical User Interface (GUI)

The WaterCare system is user friendly with a graphical interface as shown in **fig. 3** that allows user of the platform to easily use the system. When you open the app located on the local server, the clean login page appears indicating to you how to log in exactly. The interface is made to be in touch with the actual aim of the project by a background picture of actual water sample activity.

The fields to enter username and password are clearly shown to perform secure logging in. There are also different applications to service providers and new users in order to access the aspects of the system that they require in a quick manner. The interface is user-friendly and navigation among pages is simple which indicates that the front end is intact with the server. It is all about simplicity in the design; hence, it becomes easy to log in and sign up in addition to various features.

Such a well-organized GUI is what makes the system appear very professional and user-friendly so that people could easily gauge water quality.

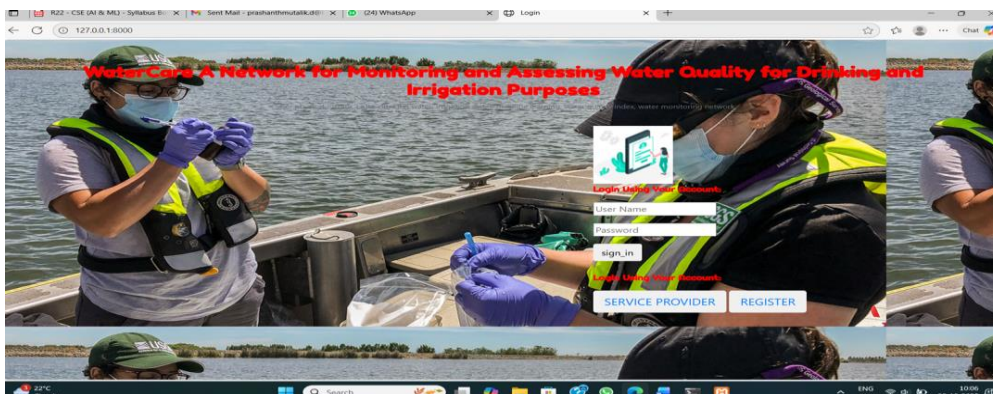


fig. 3: Graphical User Interface

4.3 Results of Remote User Management and Administrative View

The Service Provider interface has a special section, in which the admins could view all the registered remote users as shown in **fig. 4**, therefore, they could monitor the system.

As indicated in the list, WaterCare will be able to draw out user records within the database and present it in an understandable table. The rows are significant: name, email, gender, address, phone, country, state, city, which proves the fact the system stores complete registration information.

The validity of the information in the table is that the data queries are functioning correctly and user management is adequately linked to the database.

Its layout is also uncluttered and easy to navigate and view user profiles, which means that the page has a simple header.

The upper menu contains a set of options, thus admins can easily navigate data manipulation, prediction reports, accuracy outcomes and user management.

Concisely, the Remote Users page indicates that the system is operating smoothly, the data is consistently accessible and admins are in a good operation over WaterCare.

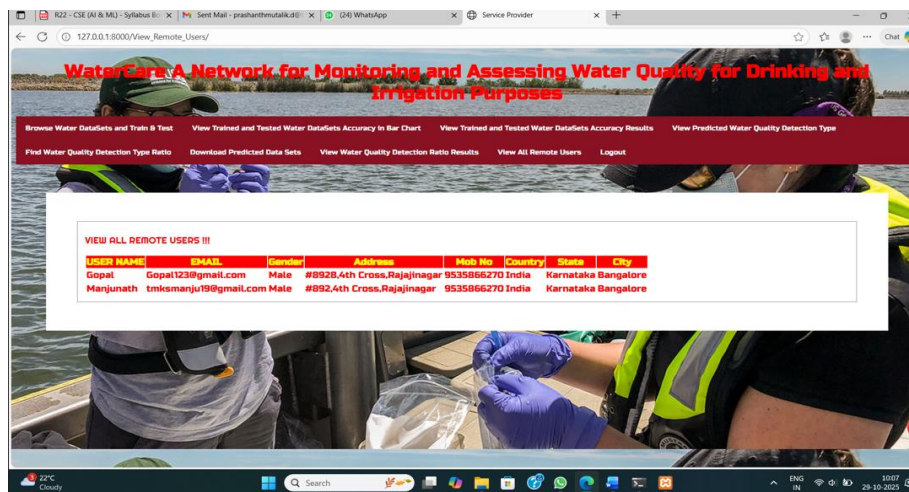


fig. 4: Admin view of Remote users

4.4 Results of Water Quality Detection Ratio Analysis

The WaterCare system presents a graph that will indicate the number of samples that consider water how many fall into the category. According to the chart, 60 percent of the samples can be used to irrigate and 40 percent as shown in **fig. 5** to drink thus the majority of the water can be used in farming. The chart is generated automatically based on the predictions that are stored within the system, which makes the results of the system to be correct and easy to analyze. The users should be allowed to change the form of charts, such as pie charts or bar graphs in order to see the data differently. Axes and percentage labels are easily read to get a person to understand how their water-quality results are distributed. On the whole, the chart helps to summarize the patterns of predictions and provides the basis of data-driven knowledge and improved decision-making in terms of water-resource assessment.

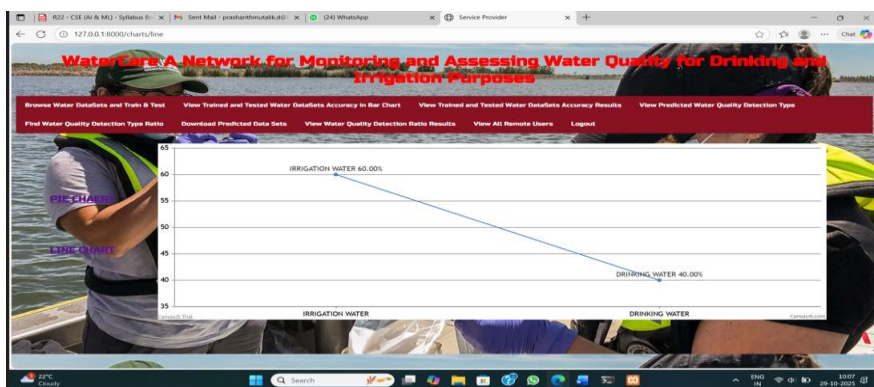


fig. 5: Water quality ratio

4.5 Results of Trained and Tested Water Dataset Accuracy Visualization

The WaterCare platform has a visual tool that the Service Provider can look at to view the performance of the trained and tested datasets. The pie chart in the interface indicates the number of samples that the machine learning models make predictions after completion of assessment of irrigation water or drinking water as shown in **fig. 6**. In this outcome, 60 1/2 of the samples are denoted as irrigation water and 40 1/2 as the drinking water providing an easy view of how the predictions are distributed between the two.

The visual is prepared on the basis of stored accuracy data and demonstrates that the system is able to receive, process and present results of prediction in a straightforward manner. The chart allows the administrators to have a quick look at the effectiveness with which the separator identifies the two types of water. The information is uncomplicated to even the non-technical people due to clear labels and the percentage values.

In general, this plot demonstrates that the classification models are reliable and that interface is presented with the results of performance in a simplified user-friendly manner to assess the effectiveness of the system.

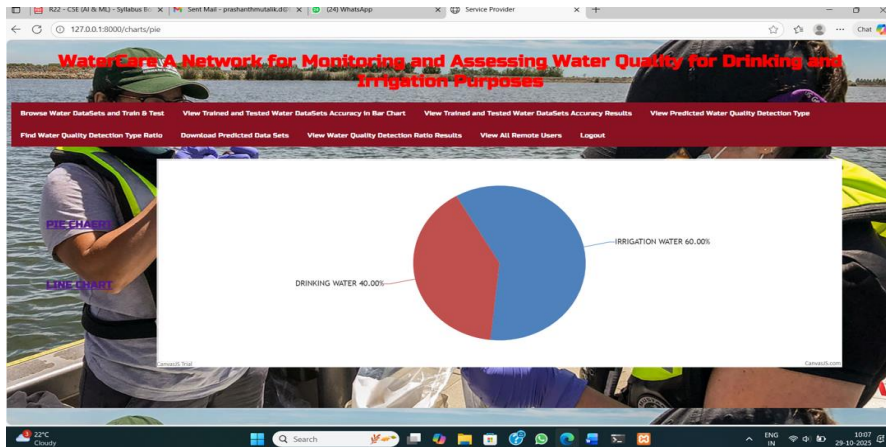


fig. 6: Visualization of water quality

4.6 Results of Water Quality Prediction for Drinking or Irrigation Use

WaterCare gives its users the chance to guess on whether or not water is safe to drink or it is merely capable of watering the plants. There is also the opportunity to enter the pH, solids, sulfates, organic carbon, hardness, chloramines, conductivity, trihalomethanes, the district, state name, and place name as shown in **fig. 7**.

Once you have entered the figures, the system executes its models immediately and places the output in the bottom.

Here the model indicates the Irrigation Water so the figures indicate that it is unsafe to drink.

The simple design, visible input fields, and fast response are an indication that the engine is functional.

The comparison with what you typed against what is displayed show that the system is reading the data properly and the model is utilizing it accordingly.

In general, the tool is an easy, quick, and convenient method of monitoring the quality of water at present.



fig. 7: Predicting Whether the water is suitable for drinking or agriculture

Discussion

The system did well with the data of water-quality and was able to provide reliable predictions on whether a sample may be utilized as a drinking or irrigation sample. Such findings demonstrate that on cleaning up the data, classifying it and presenting it through easy to relate visuals, it becomes easier to get into and utilize. The system allows people to make more effective decisions related to water-resources because real-time predictions, interactive responses, and easy-to-read output make the system user-friendly.

Nevertheless, there were some drawbacks experienced during its use by users. The performance of prediction reduces when the users provide incomplete input data or when the quality of the dataset is different. The system should also have more explicit assistance upon falling out of the expected range of the parameter in order to have users comprehend. The credibility of this system primarily hinges on the availability of all (physicochemical) data that is correct and complete.

More variety of datasets, multiple languages, and advanced result interpretation tools can be added to the system in the future, making it more adaptable, accommodating, and applicable to a wide range of users.

V.CONCLUSION

WaterCare demonstrates the effectiveness of a data-driven approach with minimal complexity in determining whether water can be used to drink or irrigate crops. It also purifies the data and employs supervised machine learning. The system examines vital physical and chemical properties and evaluates comparison of the various prediction models. It leads to the clear and reliable information which aids the users in making good decisions. The predictions can be seen by the users and managers, the analysis can be followed and the performance can be reviewed on a single screen with ease.

The system also notifies users of the features that are of most importance hence they are aware of what affects the final judgement. This provides a more acceptable view of the environment. Overall, WaterCare is a useful, expandable and user-friendly tool that allows simplifying water-quality control and gives real benefit to scientists, farmers, and the managers of common resources.

REFERENCES

- [1] Uddin MG, Nash S, Olbert AI (2021) A review of water quality index models and their use for assessing surface water quality. *Ecological Indicators* 122:107218
- [2] Foschi J, Turolla A, Antonelli M (2021) Soft sensor predictor of E. coli concentration based on conventional monitoring parameters for wastewater disinfection control. *Water Research* 191:116806
- [3] Akhtar N, Ishak MIS, Ahmad MI, Umar K, Md Yusuff MS, Anees MT, Qadir A, Almanasir YKA (2021) Modification of the water quality index (WQI) process for simple calculation using the multi-criteria decision-making (MCDM) method: A review. *Water* 13(7):905
- [4] Ortiz FM, de Almeida TT, Ferreira AE, Costa LHMK (2020) Experimental vs simulation analysis of LoRa for vehicular communications. *Computer Communications* 160:299–310
- [5] Ho L, Alonso A, Forio MAE, Vanclooster M, Goethals PLM (2020) Water research in support of the sustainable development goal 6: A case study in Belgium. *Journal of Cleaner Production* 277:124082
- [6] Banda T, Kumarasamy M (2020) Development of a universal water quality index (UWQI) for South African river catchments. *Water* 12(6):1534
- [7] Ma K, Bagula A, Nyirenda C, Ajayi O (2019) An IoT-based fog computing model. *Sensors* 19(12):2783
- [8] Raza U, Kulkarni P, Sooriyabandara M (2017) Low power wide area networks: An overview. *IEEE Communications Surveys & Tutorials* 19(2):855–873
- [9] Chilundo M, Kelderman P, O’Keeffe JH (2008) Design of a water quality monitoring network for the Limpopo river basin in Mozambique. *Physics and Chemistry of the Earth* 33:655–665
- [10] Clesceri LS, Greenberg AE, Eaton AD (2005) Standard methods for the examination of water and wastewater. American Public Health Association, Washington DC
- [11] Lee BX, Kjaerulf F, Turner S et al (2016) Transforming our world: Implementing the 2030 agenda through sustainable development goal indicators. *Journal of Public Health Policy* 37:13–31
- [12] World Health Organization (1993) Guidelines for Drinking-Water Quality. WHO Press

Copyright & License:

© Authors retain the copyright of this article. This work is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.