iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: III Month of publication: March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78870

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

LoRa- Based Dynamic Water Level and Quality Monitoring System Using Machine Learning

Devanandhini B[1], Abhirami A S[2], Anishma A R[3], Carolin Suresh[4], Prof. Jayaresmi J[5], Prof. Smitha K S[6]
[1, 2, 3, 4]BTech Student, Department of Electronics and Communication Engineering , LBS Institute of Technology for Women, Poojappura, Trivandrum, Kerala, India
[5, 6]Assistant Professor, Department of Electronics and Communication Engineering, LBS Institute of Technology for Women, Poojappura, Trivandrum, Kerala, India

Abstract: This paper proposes LoRa based Dynamic water level and quality monitoring system using machine learning. Traditional systems rely on GSM or Wi-Fi, which limits range and drains power, especially in remote areas. Recent studies show a shift toward smarter sensor networks enhanced by machine learning, but many setups still struggle with scalability, prediction, and combined monitoring of both water level and quality. The system discussed here tackles these gaps by using affordable sensors, an ESP32, and LoRa for long-range, low-power data transfer. Algorithms like Random Forest and Isolation Forest improve water quality classification and anomaly detection. Overall, combining LoRa with machine learning provides a reliable, efficient, and cost-effective way to monitor groundwater in real time and prevent contamination.
Keywords: LoRa communication, ESP32, water level, water quality, pH, TDS, turbidity, Random Forest, Isolation Forest, anomaly detection, groundwater management, low-cost IoT system.

I. INTRODUCTION

Groundwater is an important resource for homes, farms and industries especially in areas where there is not much surface water. More people moving to cities factories releasing bad stuff and changes in the weather are causing big problems like groundwater running out and getting dirty. We need to keep an eye on the water level and quality all the time to manage it well and catch any pollution early. Traditional systems that monitor water like Digital Water Level Recorders mostly just measure the water level. Use mobile phone networks to send data. These systems are expensive use a lot of energy and don't work well in areas with poor network coverage. Also checking water quality usually involves taking samples and sending them to a lab, which takes time and can't show changes in time. New technologies like the Internet of Things and wireless communication have made it possible to automate monitoring. However, many existing systems rely on Wi-Fi or mobile phone networks, which limits their coverage and increases energy use. Most systems also just rely on rules to analyze data and don't have smart decision-making capabilities.

This paper suggests a groundwater monitoring system that uses cheap sensors long-range communication and machine learning. The system monitors water level and key quality parameters like pH, turbidity and Total Dissolved Solids in time. It uses a Random Forest classifier to classify water quality and an Isolation Forest algorithm to detect anomalies. It also has an energy- way of sending data to save power.

The proposed system offers a affordable and smart solution for monitoring groundwater, especially suitable for remote and resource-limited environments.

It can help us manage groundwater better and catch any problems early.

The system is an improvement over traditional systems and can make a big difference, in areas where groundwater is crucial.

II. LITERATURE REVIEW

In [1] Kombo, Omar H., Santhi Kumaran, and Alastair Bovim proposed a low-cost, low-power LoRa-GSM IoT-enabled system for monitoring groundwater resources with energy harvesting integration. The design uses an Arduino UNO and solar power for continuous water level measurement, transmitting data several times daily via LoRa-GSM hybrid communication. This system effectively demonstrated the feasibility of combining renewable energy and IoT for remote monitoring. However, it focuses only on water level measurement, lacking advanced machine learning analytics, multi-parameter water quality assessment, and a predictive dashboard for real-time interpretation.

In [2] Rahman, Md. Mahbubur, Chinmay Bapery, Mohammad Jamal Hossain, Zahid Hassan, G. M. Jamil Hossain, and Md. Muzahidul Islam introduced an IoT-based water quality monitoring system using Arduino and Node MCU. Sensor data was transmitted via Wi-Fi to a Firebase Realtime Database and displayed on a web interface for visualization. The framework successfully demonstrated real-time monitoring of pH, turbidity, and temperature. However, the system suffered from the limited range of Wi-Fi communication and high-power consumption, making it impractical for large-scale or remote applications where internet connectivity is unavailable.

In [3] Fei Tony Liu, Kai Ming Ting, and Zhi-Hua Zhou presented the Isolation Forest algorithm for anomaly detection, which isolates anomalies using random partitioning. This method identifies unusual data points based on shorter path lengths in randomly generated isolation trees. The approach became a significant basis for unsupervised anomaly detection in environmental data and sensor networks. Despite its efficiency, the Isolation Forest is highly sensitive to parameters like subsample size and number of trees, and although it produces anomaly scores, it offers limited interpretability regarding the causes of detected anomalies.

In [4] a Water Quality Monitoring System using a Machine Learning Model was discussed in the International Journal of Creative Research Thoughts (IJCRT, 2024).

The authors employed the Random Forest Classifier (RFC) for real-time water quality monitoring visualization and reporting. However, sensor dependency and noise interference affected reliability, and the system remained at the prototype stage, without large- scale field validation or energy optimization.

In [5] Andrew J. Calderwood, Richard A. Pauloo, Alysa M. Yoder, and Graham E. Fogg developed a low-cost, open-source wireless sensor network for real-time groundwater monitoring using LoRa communication. The system, known as the Groundwater Observatory, utilized pressure transducers and a cellular telemetry network for data collection and visualization through SQLite/R/Shiny tools. While the system provided scalable and open-source monitoring capabilities, it required reliable cellular connectivity for data transmission. Moreover, periodic manual calibration of sensors was necessary, reducing long-term autonomy in remote areas.

In [6] Zehua Sun, Huanqi Yang, Kai Liu, Zhimeng Yin, Zhenjiang Li, and Weitao Xu presented a comprehensive survey on LoRa technology covering its performance, communication structure, security aspects, and emerging IoT applications. The survey analysed LoRa using analytical models, simulators, and experimental test beds. Although this paper provides extensive insight into LoRa's technical capabilities and potential applications, it is purely theoretical and lacks experimental validation or implementation of real-world monitoring systems.

In [7] Erianto Indra Putra, Muhammad Uqbah El Syakbandani, Sigit Pramono, and Asmadi Saad investigated the use of submersible sensors to estimate groundwater level in peatlands. The study compared data from automated sensors against manual measurements across multiple sites and used statistical t-tests to assess accuracy. The results showed that submersible sensors produced reliable results with discrepancies ranging from 0.1–1 cm, attributed mainly to sensor drift and calibration errors. However, dependence on manual validation limited scalability, and maintenance issues made it challenging for large or remote deployments.

III. PROPOSED SYSTEM

The system we are talking about has three parts: sensor nodes, a communication module and a data processing unit.

The sensor nodes have a lot of low-cost sensors that are connected to an ESP32 microcontroller. These sensors measure things like the water level, how acidic the water is, how cloudy it is and how much stuff is dissolved in it. The sensor nodes collect data all the time. It is real-time data about the environment. This data is sent using LoRa communication modules. LoRa is good because it uses little power and can send data a long way without using the internet or phone networks. This means we can use the system in places that're really far away, from everything.

When the data gets to where it's going it is looked at and analyzed using machine learning models. We made a dashboard using Flask that shows the real-time data, what the system thinks the data means and if anything, weird is happening. The system shows water level data, pH data, turbidity data and TDS data in time.

IV. METHODOLOGY

The methodology consists of four main stages: data collection, preprocessing, model training, and data visualization. The figure shows the basic block diagram of the system.

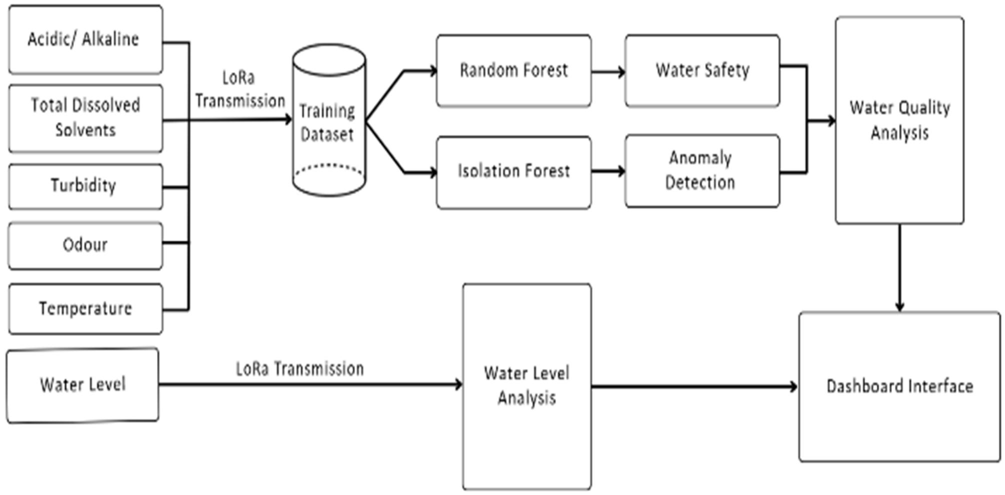


Figure 1: Block Diagram

The important steps are described below:

1) Data collection and Preprocessing: The dataset for this project is generated through real-time sensor measurements collected from various water quality and level monitoring nodes. Each node is equipped with sensors for measuring pH (acidic/alkaline nature), total dissolved solvents (TDS), turbidity, odour, and water pressure. The data from these sensors are transmitted via LoRa (Long Range) wireless communication to ESP32 microcontroller. Upon collection, the raw data undergoes preprocessing to ensure consistency and reliability. This involves removing outliers, handling missing values, and normalizing sensor readings across different units and ranges. These preprocessing steps eliminate noise and discrepancies caused by environmental variations, sensor drift, or temporary communication errors, ensuring that the dataset accurately reflects water conditions.
2) Dataset Formation and Feature Engineering: After preprocessing, the cleaned readings are compiled into a structured training dataset that serves as the foundation for model training. Additional features such as average daily values, rate of change in turbidity, and pH deviation trends are derived to improve the model's learning capability. The dataset is then labeled based on standard water safety guidelines (e.g., safe, moderate, or unsafe) to facilitate supervised learning. This structured and feature-enhanced dataset helps the models better capture correlations among water parameters and predict quality and anomalies more effectively.
3) Machine Learning Model Training: Two primary machine learning algorithms are employed in this system:

a) Random Forest Classifier: This algorithm is trained using the labeled dataset to evaluate water safety levels. By aggregating decisions from multiple decision trees, the Random Forest model effectively classifies water as safe, moderate, or unsafe, based on sensor data patterns.
b) Isolation Forest Algorithm: This unsupervised learning model is used for anomaly detection. It isolates unusual data points that may indicate contamination, sensor faults, or abnormal changes in water quality. Together, these models enable both predictive analysis and early detection of irregularities, providing a comprehensive water monitoring solution

4) Water level analysis and Validation: In parallel with quality monitoring, water pressure readings are analyzed to determine water levels. The data is transmitted via LoRa and processed to calculate height variations using calibrated pressure-to-depth relationships.
5) Dashboard Interface Development: A Flask-based web dashboard is developed to provide real-time visualization of water quality and level data. It serves as an interactive interface where users can monitor parameters such as pH, TDS, turbidity, odor, and water level through live charts and status indicators. The dashboard retrieves processed results from the backend machine learning models and displays water safety status and anomaly alerts in a clear, user-friendly layout. This integration of Flask ensures a lightweight, responsive, and easily accessible monitoring platform for continuous supervision of remote water sources.

V. MACHINE LEARNING FRAMEWORK

This stage is the analytical core of the system, applying machine learning algorithms for intelligent decision-making.

1) Random Forest Classifier: The model categorizes water quality as safe, moderate, or unsafe based on parameters like pH, TDS, and turbidity. By learning from historical sensor data, it provides a more adaptive and accurate classification than conventional threshold-based methods.
2) Isolation Forest Algorithm: This algorithm identifies anomalies in the data that may signify contamination, equipment malfunction, or abrupt environmental changes.

Together, these models provide reliable and autonomous detection of quality deterioration and abnormal water level variations, supporting early warning and preventive action.

VI. SYSTEM REQUIREMENTS

A. Software Requirements

1) Visual Studio Code (VS Code):Visual Studio Code (VS Code) is the primary Integrated Development Environment (IDE) used in this project. VS Code is an open-source, lightweight, and highly customizable IDE that supports Python programming and offers a range of extensions ideal for machine learning and web development.
2) Jupyter Notebook / Google Colab: Jupyter Notebook and Google Colab were used for developing, testing, and training the machine learning models. These platforms provide an interactive environment where code, visualizations, and documentation can coexist. Google Colab provides free GPU support and cloud storage, facilitating faster model training and sharing of experiments without local computational limitations.
3) Python Software (Version 3.9.13): Python is the primary language for developing this project's backend. It provides an extensive ecosystem of libraries necessary for machine learning, data manipulation, and web development.
4) Flask (Python Web Framework): Flask is a lightweight and flexible web framework used to create the interactive web dashboard for this project. It enables seamless integration of the trained machine learning models with a user-friendly interface where sensor data can be input, and results such as Safe or Unsafe water classification are displayed in real time. Flask's simplicity allows rapid development and easy deployment of web applications.
5) Scikit-learn Library: The Scikit-learn library was utilized to implement the core machine learning algorithms for the project, including the Random Forest Classifier, Isolation Forest for anomaly detection, and Linear Regression for predictive analysis.
6) NumPy & Pandas Libraries: NumPy and Pandas were used for handling numerical computations and managing sensor data efficiently. NumPy provides an optimized array of operations and mathematical functions, while Pandas enables structured data manipulation, cleaning, and analysis.
7) Joblib / Pickle: Joblib and Pickle were used to save and load the trained machine learning models. After training, the models were serialized and stored, enabling the Flask application to load them efficiently without retraining them. This ensures faster response times and smooth integration between the backend model computations and the web dashboard.

B. Hardware Requirements

The Machine Learning–based Intelligent Water Quality Monitoring System relies on a combination of sensors and microcontrollers to capture accurate water quality parameters in real time. The hardware components ensure precise data acquisition, reliable communication, and seamless integration with the software for processing and visualization

1) LoRa Module: The LoRa (Long Range) module enables wireless transmission of sensor data over long distances with minimal power consumption. This allows the water quality monitoring system to transmit data from remote or distributed locations to a central server or dashboard.
2) ESP32 Microcontroller: The ESP32 is a powerful, low-cost microcontroller with built-in Wi-Fi and Bluetooth connectivity. It serves as the central processing unit for collecting sensor data, performing preliminary data processing, and transmitting information to the Flask web dashboard
3) pH sensor: The pH sensor measures the acidity or alkalinity of water. Maintaining proper pH levels is critical for safe drinking water and environmental health. The sensor provides analog or digital readings that are processed by the ESP32 and fed into the machine learning model for water quality classification.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

4) Turbidity Sensor: The turbidity sensor detects the cloudiness or suspended particle content in water. High turbidity indicates contamination and poor water quality. The sensor readings are used by the system to determine if the water meets safety thresholds, triggering alerts when turbidity exceeds acceptable limits.
5) Total Dissolved Solids (TDS) sensor: The TDS sensor measures the concentration of dissolved salts and minerals in water. High TDS levels can indicate contamination or salinity issues. These readings are critical for evaluating overall water quality and are integrated into the machine learning classification algorithm.
6) Ammonia Gas Sensor: The MQ sensor is used to detect gases emitted from the water surface, which may indicate the presence of contamination such as organic decomposition or chemical reactions producing volatile compounds.
7) Ultrasonic Sensor: The ultrasonic sensor measures the water level that helps in drought prediction.
8) Temperature Sensor: The temperature *sensor measures the temperature of water, as temperature affects the physical and chemical properties of water.*

VII. EXPERIMENTAL RESULTS AND DISCUSSION

A. *Experimental Setup*

The proposed system was implemented using an ESP32 microcontroller interfaced with multiple sensors including pH, turbidity, TDS, MQ gas sensor, temperature sensor, and ultrasonic sensor for water level measurement. The hardware setup consists of a transmitting unit and a receiving unit connected via LoRa communication.

The transmitting unit collects real-time sensor data and displays it locally on an LCD screen. The collected data is transmitted wirelessly using LoRa modules to the receiver, where it is processed and visualized on a Flask-based dashboard.

The system was tested under different water conditions to validate sensor accuracy, communication reliability, and machine learning performance.

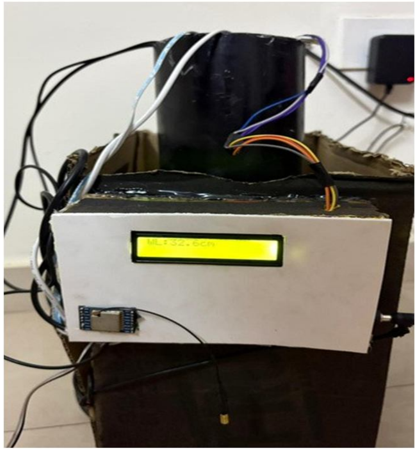


Figure 2: Transmission Side Setup

1) *Water Quality Standards:* To evaluate water safety, standard permissible limits were considered based on drinking water quality guidelines. These thresholds were used to label the dataset and validate system performance.

S. No.	Parameters	Permissible Value	Standard
1	Colour	Unobjectionable	IS: 10500
2	Taste	Agreeable	IS: 10500
3	pH	6.5 – 7.5	IS: 10500
4	Turbidity	5	IS: 10500
5	TDS	500	IS: 10500
6	TSS	5	USPHS
7	BOD	Nil – 5	USPHS
8	DO	4 – 6	USPHS
9	Total Hardness	300	IS: 10500
10	Chloride	250	IS: 10500
11	Alkalinity	120	USPHS
12	Residual Chlorine	0.2	IS: 10500

Table 1: Water Quality Standards

When any of these parameters went beyond the safe range for example, turbidity exceeding 5 NTU, pH dropping below 7 or rising above 8.5, or temperature beyond the 20–30°C range the system automatically classified the sample as "Unsafe" and highlighted it on the dashboard. In addition, the system detected sudden spikes in any of the parameters, marking them as potential anomalies and triggering a warning message. This ensured early detection of pollution, contamination, or equipment faults.

2) Good Water Quality

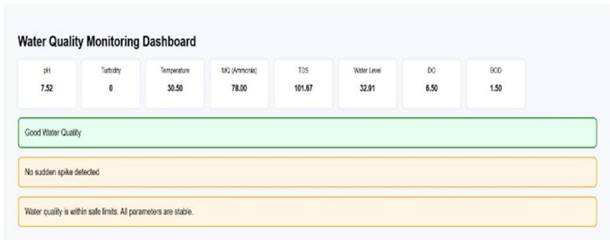


Figure 3: Good Water Quality Output

The figure shows the dashboard output when all monitored parameters are within their permissible limits. Parameters such as pH, turbidity, temperature, and TDS remain stable, resulting in the classification of "Good Water Quality." The system also confirms "No sudden spike detected," indicating the absence of anomalies and ensuring that the water is safe for usage.

Figure 4: Good Water Quality Analytics

The figure shows the sensor analytics for good water quality conditions. All parameters exhibit minimal variation and maintain consistent trends over time. This stability confirms the reliability of sensor measurements and demonstrates the system's ability to maintain accurate monitoring under normal environmental conditions.

3) Moderate Water Quality

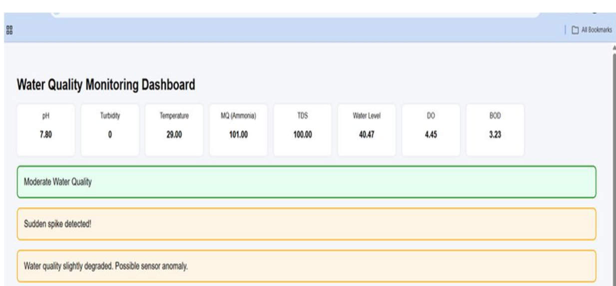


Figure 5: Moderate Water Quality Output

The figure depicts the dashboard output under moderate water quality conditions. Slight deviations in parameters such as TDS, temperature, or pH are observed, leading to the classification of "Moderate Water Quality." The system detects minor anomalies, indicating early signs of water quality degradation that may require monitoring or preventive action.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

Figure 6: Moderate Water Quality Analytics

This shows the analytics graph corresponding to moderate water quality. Gradual fluctuations in certain parameters are visible, suggesting deviations from normal conditions. These trends highlight the system's sensitivity in detecting early-stage variations before they escalate into critical contamination.

4) Poor Water Quality

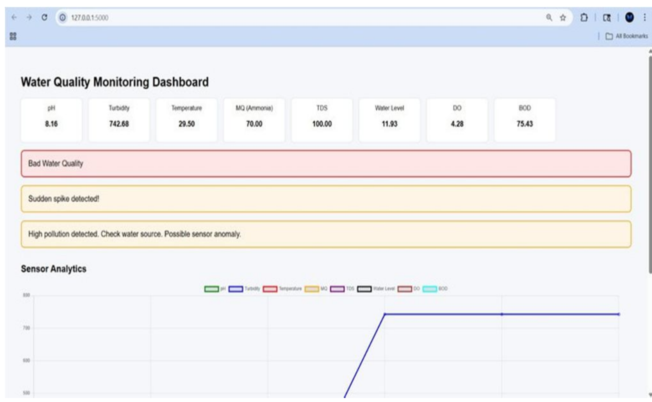


Figure 7: Poor Water Quality Output

The figure presents the dashboard output during poor water quality conditions. Multiple parameters exceed safe limits, leading to the classification of "Bad Water Quality." The system generates alerts such as sudden spike detection and possible contamination warnings, indicating that the water is unsafe and requires immediate attention.

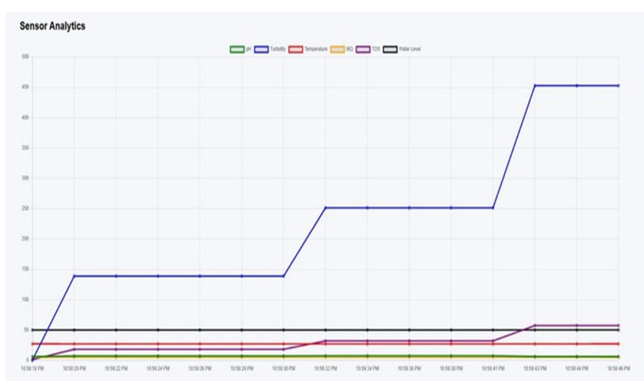


Figure 8: Poor Water Quality Analytics

This figure presents the dashboard output during poor water quality conditions. Multiple parameters exceed safe limits, leading to the classification of "Bad Water Quality." The system generates alerts such as sudden spike detection and possible contamination warnings, indicating that the water is unsafe and requires immediate attention.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

5) Drought Prediction

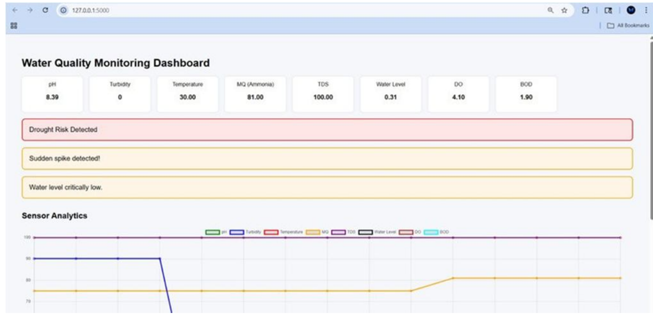


Figure 9: Drought Prediction Output

The system detects a significant drop in water level and triggers a “Drought Risk Detected” alert. Additional warnings such as sudden spike detection and critically low water level indicate abnormal environmental conditions requiring immediate attention.

Figure 10: Drought Prediction Analytics

A sharp decline in water level is observed, while other parameters remain relatively stable. This confirms that the system effectively identifies drought scenarios based on water level variations.

6) Machine Learning Model Evaluation

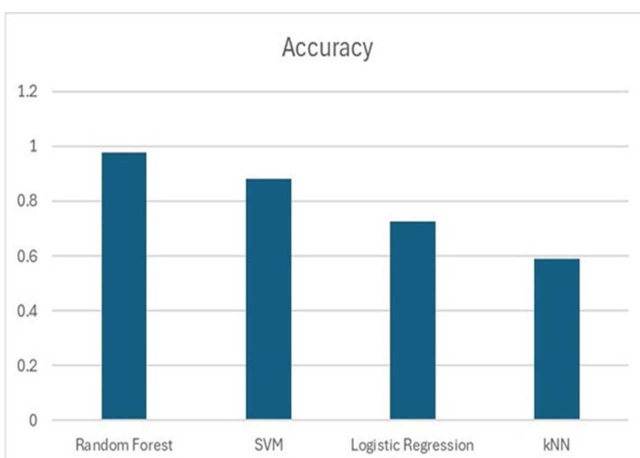


Figure 11: Model Accuracy Chart

The chart shows the comparison of classification accuracy among different machine learning models. The Random Forest model achieves the highest accuracy compared to SVM, Logistic Regression, and kNN, demonstrating its effectiveness for water quality prediction.

Scores

Model	AUC	CA	Prec
Random Forest	0.999	0.973	0.973
SVM	0.968	0.880	0.884
Logistic Regression	0.819	0.725	0.732
kNN	0.638	0.590	0.577

Figure 12: Comparison Table

The figure presents the performance comparison of different machine learning models based on Area Under Curve (AUC), classification accuracy (CA), and precision. The Random Forest model outperforms other models with the highest AUC and accuracy, making it the most suitable choice for the proposed system.

VIII. CONCLUSION

The Machine Learning–based Intelligent Water Quality Monitoring System represents a significant advancement in integrating IoT and artificial intelligence for environmental monitoring. Now we have successfully developed and deployed a Flask-based web dashboard capable of real-time monitoring and classification of water quality based on five key parameters: pH, turbidity, temperature, dissolved oxygen (DO), and conductivity. The system accurately classifies water as “Safe” or “Unsafe”, detects sudden spikes and anomalies in sensor readings using Isolation Forest, and predicts future trends using Linear Regression. This phase has validated the feasibility of combining machine learning models with IoT sensors to provide actionable insights into water quality. Next phases will expand the project toward a comprehensive smart water management solution. The ongoing development promises a scalable, intelligent, and sustainable approach to ensuring safe water resources for communities

REFERENCES

[1] O. H. Kombo, S. Kumaran, and A. Bovim, “Design and Application of a Low-Cost, Low- Power, LoRa-GSM, IoT Enabled System for Monitoring of Groundwater Resources with Energy Harvesting Integration,” IEEE Access, vol. 9, pp. 150312–150324, 2021.

[2] M. M. Rahman, C. Bapery, M. J. Hossain, Z. Hassan, G. M. J. Hossain, and M. M. Islam, “Internet of Things (IoT) Based Water Quality Monitoring System,” International Journal of Multidisciplinary and Current Educational Research (IJMCER), vol. 2, no. 4, pp. 245–252, 2020.

[3] F. T. Liu, K. M. Ting, and Z. H. Zhou, “Isolation Forest,” Proceedings of the IEEE International Conference on Data Mining (ICDM), Pisa, Italy, pp. 413–422, 2008.

[4] “Water Quality Monitoring System Using ML Model,” International Journal of Creative Research Thoughts (IJCRT), vol.12, no. 5, pp. 1124–1130, 2024.

[5] A. J. Calderwood, R. A. Pauloo, A. M. Yoder, and G. E. Fogg, “Low-Cost, Open-Source Wireless Sensor Network for Real-Time, Scalable Groundwater Monitoring,” Water, vol. 12, no. 4, pp. 1–16, 2020.

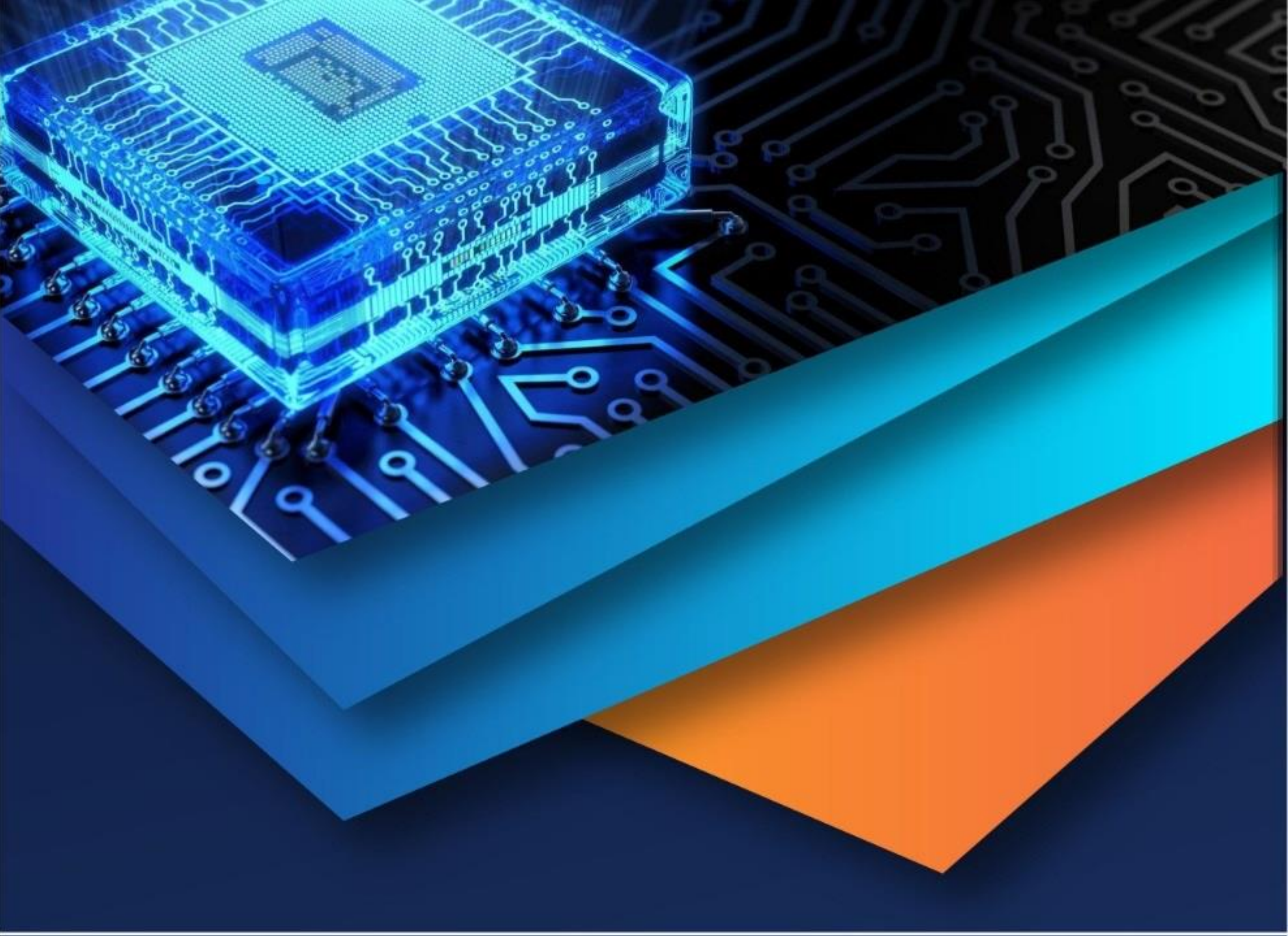
[6] Z. Sun, H. Yang, K. Liu, Z. Yin, Z. Li, and W. Xu, “Recent Advances MS56in LoRa: A Comprehensive Survey,” ACM Transactions on Sensor Networks (TOSN), vol. 18, no. 4, pp. 1–39, 2022.

[7] E. I. Putra, M. U. E. Syakbandani, S. Pramono, and A. Saad, “Estimating Groundwater Level in Peatlands by Using Submersible Sensor,” Journal of Tropical Silviculture Science and Technology, vol. 15, no. 1, pp. 25–32, 2024.

[8] V. Shilwant, P. Thakare, and S. Sangle, “Comparative Study of Water Quality Assessment with IoT & Lab Test,” International Journal of Engineering Research & Technology (IJERT), vol. 13, no. 4, pp. 450–453, 2024.

[9] G. R. Kumar, D. Kishore, and G. V. M. G. Kumar, “Waste Contamination in Water – A Real-Time Water Quality Monitoring System using IoT,” Proceedings of the International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, IEEE, pp. 1–5, 2021.

[10] A. N. Prasad, K. A. Mamun, F. R. Islam, and H. Haqva, “Smart Water Quality Monitoring System,” Proceedings of the Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), Nadi, Fiji, IEEE, pp. 1–6, 2015.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY