



AI-Powered Water Quality Detection and Purification Recommendation System Using Refactor Index

Abin Santhosh¹, Adhith Sunil², Adhwaith T T³, Anitta Raphi E⁴, Meethu M B⁵

^{1,2,3,4}Student, Department of Computer Science and Engineering, IES College of Engineering, Thrissur, Kerala, India

⁵Assistant Professor, Department of Computer Science and Engineering, IES College of Engineering, Thrissur, Kerala, India

Email_id: abinsanthosh19@gmail.com, suniladhith1@gmail.com, adhwaithtt03@gmail.com, anittaraphi@gmail.com, meethumb@iesce.info

Abstract

The proposed project introduces a smart water quality detection and purification recommendation system powered by artificial intelligence. The system uses sensors such as pH, turbidity, temperature, and dissolved oxygen to collect water data in real time. An ESP32 microcontroller with Wi-Fi and Bluetooth processes the readings and calculates a Refactor Index (RI), a single score that classifies water as “Good,” “Moderate,” “Poor,” or “Unsafe.” Results are displayed on an OLED screen and shared with users through a mobile application. Based on the RI score, the system also suggests suitable purification methods such as boiling, filtration, or UV treatment. This solution is low-cost, portable, and eco-friendly, making it useful for households, rural communities, aquaculture, and agricultural applications. By integrating AI and IoT technologies, the project ensures real-time monitoring, reduces health risks, and promotes sustainable water management.

Keywords: Water Quality Detection, Refactor Index, IoT Sensors, ESP32 Microcontroller, Artificial Intelligence, Real-Time Monitoring, Environmental Sustainability, Smart Water System.

DOI: <https://doi.org/10.5281/zenodo.18831716>

1. Introduction

In today’s rapidly evolving world, access to clean and safe water remains a critical global concern. Traditional water quality assessment methods such as laboratory-based chemical and biological testing while precise, are often time-intensive, costly, and unsuitable for real-time monitoring. These limitations hinder timely detection of contamination, especially in rural and resource-constrained environments, increasing the risk of waterborne diseases and undermining public trust in water sources. Industries, agriculture, and aquaculture also face operational challenges due to inconsistent water evaluations and delayed interventions.

To overcome these barriers, this project proposes an AI-powered Water Quality Detection and Purification Recommendation System that combines artificial intelligence with Internet of Things (IoT) technologies to deliver a smart, automated, and scalable solution. The system utilizes sensors—including pH, turbidity, temperature, and dissolved oxygen—connected to an ESP32-WROOM-32 microcontroller to collect and process water data in real



time. A composite metric, the Refactor Index (RI), is calculated to classify water quality into four categories: Good, Moderate, Poor, or Unsafe. Based on the RI score, the system recommends suitable purification methods such as boiling, filtration, or UV treatment.

1.1. Sensor-Based Data Acquisition:

Real-time water parameters are captured using low-cost sensors integrated with the ESP32 microcontroller, enabling continuous monitoring across diverse environments.

1.2. Refactor Index (RI) Computation:

A composite scoring algorithm processes sensor inputs to generate a simplified water quality index, facilitating intuitive classification and decision-making.

1.3. AI-Driven Purification Recommendation Module – Based on the RI score, the system intelligently suggests appropriate purification techniques tailored to the detected contamination level.

1.4. User Interface and Connectivity – Results are displayed on an OLED screen and transmitted via Wi-Fi/Bluetooth to a mobile application, offering users instant feedback and actionable insights.

1.5. Sustainable and Scalable Deployment – Powered by solar energy, the system is designed for portability and eco-friendliness, making it ideal for households, rural communities, and industrial applications.

2. Literature Survey

The paper by **Lili Jin, Hui Huang, and Hongqiang Ren** [1] explores the transformative potential of artificial intelligence in optimizing water treatment technologies. Their study emphasizes how AI can automate complex purification workflows, reduce operational costs, and enhance predictive maintenance in industrial water systems. Although focused on large-scale applications, the paper lays a foundational understanding of how intelligent algorithms can be adapted for smaller, portable systems like the one proposed in this project. It validates the use of AI not just for detection but also for recommending context-aware purification strategies.

Preeti Verma and Pankaj Mehta [2] present a review of emerging trends in real-time water quality monitoring using IoT-enabled sensor networks. Their work highlights the global sanitation challenges and the importance of low cost, scalable solutions. The integration of sensors with microcontrollers for continuous data acquisition aligns directly with your system's architecture. However, their study lacks a unified scoring mechanism or decision-support module gaps that your Refactor Index and AI-driven purification recommendations aim to fill.

Esra Kendir and Şerafettin [3] investigate the effect of temperature and wavelength on the refractive index of water using fiber-optic sensors. Their experimental findings support the use of refractive index as a viable parameter for water purity analysis. This directly reinforces your project's innovation in incorporating RI as a composite metric, offering a novel approach to water classification beyond traditional chemical parameters.

Yuanfeng Qi and Kai He [4] survey advanced water purification technologies, including membrane filtration, adsorption, and hybrid oxidation systems. While their taxonomy is comprehensive, it lacks integration with real-time detection systems. Your project bridges this gap by not only detecting impurities but also recommending purification methods based on live sensor data and AI inference, making the process more responsive and user-centric.



Ramakant, Shuchi, and Manvi [5] discuss recent developments in hybrid oxidation systems and AI-enhanced membrane technologies. Their work demonstrates how machine learning can optimize purification efficiency, especially in industrial contexts. However, the absence of edge computing and mobile deployment limits its accessibility. Your system's use of ESP32 and mobile app integration offers a more inclusive and field-ready alternative.

Ansari et al. [6] propose an AI model for predicting multiple water quality parameters using historical datasets. Their system achieves high accuracy in forecasting but lacks real-time sensor integration. Your project advances this by embedding AI directly into edge devices, enabling on-site inference and immediate feedback—critical for rural and household applications.

Gutenson [7] introduces a rule-based expert system for water decontamination decisions. While effective in structured environments, the system lacks adaptability and scalability. Your AI model offers a more flexible and data-driven approach, capable of learning from diverse water conditions and evolving over time. The RI scoring system also simplifies user interpretation compared to rigid rule-based outputs.

Shaghghi [8] presents DOxy, a low-cost IoT system calibrated for dissolved oxygen sensing using pulse-oximetry. The study demonstrates the feasibility of affordable DO monitoring, aligning with your project's emphasis on accessibility and sensor modularity. However, DOxy focuses on a single parameter, whereas your system integrates multiple sensors for a holistic water quality assessment.

Dr. B. Shraavan Kumar et al. [9] develop an Arduino-based water quality monitoring system with pH, temperature, and water level sensors. Their system includes automation features but lacks a unified scoring mechanism like the Refactor Index. Your RI model simplifies interpretation and enhances user understanding, making water safety decisions more intuitive.

Nur Amalina Binti Rosle and Bin Alias [10] propose a turbidity monitoring system using low-cost sensors and real-time data transmission. While effective for single-parameter tracking, the system does not support multi-sensor fusion or purification recommendations. Your project's composite RI and AI module offer a more holistic solution, enabling both detection and actionable guidance.

3. Review Of Methodology

3.1 System Design:

The AI-Based Water Quality Detection and Recommendation System integrate real-time sensor data, edge computing, and cloud-based intelligence to assess water quality and suggest appropriate purification methods. Key stakeholders include users, sensor hardware, the ESP32 microcontroller, AI models, and the mobile/web interface. Sensors collect water parameters, the ESP32 processes and computes the Refactor Index (RI), and the AI module recommends purification techniques. The system ensures portability, reliability, and eco-friendly operation for diverse environments such as households, agriculture, and aquaculture.

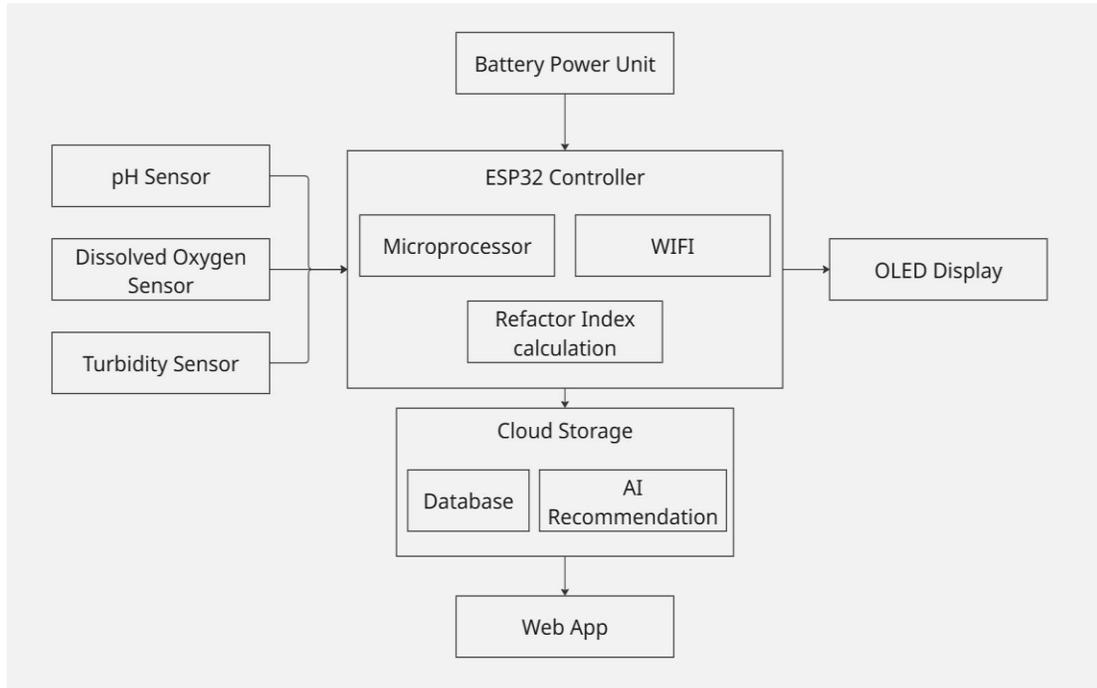


Figure 1: System Design

The AI-Based Water Quality Detection and Recommendation System integrates real-time sensor data, edge computing, and cloud-based intelligence to assess water quality and suggest appropriate purification methods. Key stakeholders include users, sensor hardware, the ESP32 microcontroller, AI models, and the mobile/web interface. Sensors collect water parameters, the ESP32 processes and computes the Refractor Index (RI), and the AI module recommends purification techniques. The system ensures portability, reliability, and eco-friendly operation for diverse environments such as households, agriculture, and aquaculture.

3.2. User Module

The User Module ensures intuitive and secure interaction, allowing users to:

- View Real-Time Data: Access water quality readings (pH, turbidity, temperature, DO) via OLED display or mobile/web interface.
- Receive RI Classification: Understand water safety levels categorized as Good, Moderate, Poor, or Unsafe.
- Get Purification Suggestions: Receive AI-driven recommendations such as boiling, filtration, or UV treatment based on RI score.
- Monitor History: View historical water quality trends and purification logs via cloud dashboard.

3.3. Data Acquisition and Preprocessing Module

This module collects and prepares sensor data for analysis:

- Sensor Integration: pH, turbidity, temperature, and dissolved oxygen sensors transmit analog/digital signals to ESP32.
- Signal Conditioning: Filters noise and calibrates raw sensor data for accuracy.



- c. Normalization: Standardizes values across different sensor ranges to enable unified RI computation.
- d. Timestamping: Associates each reading with time metadata for historical tracking and trend analysis.

3.4. Refactor Index (RI) Computation Module

This module calculates a composite score to simplify water quality interpretation:

- a. Weighted Scoring Algorithm: Assigns weights to each parameter based on health impact and environmental standards.
- b. RI Classification: Maps the computed score to predefined categories (Good, Moderate, Poor, Unsafe).
- c. Threshold Calibration: Dynamically adjusts classification thresholds based on regional water standards or user-defined preferences.
- d. Edge Inference: Performs RI computation locally on ESP32 for low-latency feedback.

3.5. AI Recommendation Module

This module suggests purification methods based on RI and sensor trends:

- a. Rule-Based Filtering: Applies decision rules to match RI scores with suitable purification techniques.
- b. Adaptive Learning: Incorporates feedback from user choices and historical outcomes to refine recommendations.
- c. Multi-Condition Handling: Suggests hybrid methods (e.g., boiling + filtration) for complex contamination profiles.
- d. Cloud Sync: Updates AI model periodically with new data for improved accuracy and personalization.

3.6. Communication and Display Module

This module ensures seamless data visualization and user interaction:

- a. OLED Display: Shows real-time RI score and purification advice on-device.
- b. Mobile/Web Interface: Provides remote access to water quality data, graphs, and recommendations.
- c. Wi-Fi/Bluetooth Sync: Enables data transmission between ESP32 and cloud/database.
- d. Alert System: Sends notifications for unsafe water conditions or sensor anomalies.

3.7. Power and Sustainability Module

This module ensures uninterrupted and eco-friendly operation:

- a. Battery Power Unit: Supplies energy to all components with rechargeable capability.
- b. Solar Integration: Optional solar panel support for off-grid deployment.
- c. Power Optimization: Implements sleep cycles and low-power modes on ESP32 to extend battery life.

4. Review Of Datasets

The AI-Based Water Quality Detection and Purification Recommendation System relies on a well-defined dataset architecture to ensure accurate water quality assessment and actionable purification guidance. The dataset is organized into four core entities—Sensors, Users, Readings, and Recommendations—each with specific attributes and relationships. This structured approach supports efficient data storage, retrieval, and analysis, enabling real-time monitoring and intelligent decision-making.

4.1. Sensor Dataset

This subset captures water quality parameters from hardware sensors. Each sensor is uniquely identified by a



SensorID and described by its type (such as pH, turbidity, temperature, or dissolved oxygen) and its measurement unit. The sensor data is collected in real time via the ESP32 microcontroller and stored with metadata to ensure traceability and accuracy. These sensors serve as the primary input sources for the system's analytical pipeline.

4.2. User Dataset

The user dataset manages participant-specific information and access control. Each user is assigned a unique UserID and associated with attributes such as name, email, and location. This dataset enables personalized tracking water quality readings and supports location-based filtering for regional analysis. By linking users to their readings, the system ensures contextual relevance and facilitates targeted feedback and notifications.

4.3. Readings Dataset

This is the core analytical subset that stores real-time water quality data. Each reading is identified by a ReadingID and linked to both a SensorID and a UserID. It includes the recorded value, timestamp, computed Refactor Index (RI) score, and the interpreted water quality status (e.g., Good, Moderate, Poor, Unsafe). This dataset forms the foundation for purification recommendations and supports historical trend analysis and AI model training.

4.4. Recommendation Dataset

The recommendation dataset provides purification guidance based on RI scores. Each entry includes a unique RecID, a reference to the corresponding ReadingID, the suggested purification method (such as boiling, filtration, UV, or chemical treatment), and a descriptive explanation. By mapping recommendations to specific readings, the system ensures that users receive actionable and context-aware advice tailored to the detected water quality.

4.5. Dataset Utility and Scalability

The overall dataset architecture is designed for scalability, security, and extensibility. Indexed fields enable fast querying for mobile and web interfaces, while historical data supports adaptive learning and predictive modeling. Role-based access and encryption safeguard sensitive information, and the schema allows for future expansion to include additional parameters like TDS, EC, or heavy metals. This robust design ensures the system remains reliable, responsive, and adaptable across diverse deployment scenarios.

5. Implementation of the Multimodal Ai-Based Mental Health Detection System

The system is built to deliver real-time water quality monitoring and purification guidance using sensor data, AI processing, and a web-based interface. It combines hardware sensing with intelligent software to ensure fast, reliable, and user-friendly operation.

5.1. Hardware Requirements

- **a. ESP32-WROOM-32** microcontroller for data processing and wireless communication
- **b. Sensors:** pH, turbidity, temperature, and dissolved oxygen
- **c. OLED Display** for showing RI scores and recommendations
- **d. battery** for operation
- **e. Connectivity** via Wi-Fi/Bluetooth
- **f. Supporting components:** breadboard, jumper wires, voltage regulator



5.2. Software Requirements

- **a. Front-End:** React with TypeScript for a responsive web interface
- **b. Back-End:** Python (Django/Flask) for RI computation and API handling
- **c. AI Integration:** TensorFlow or Scikit-learn for purification logic
- **d. Database:** SQLite for storing readings and recommendations
- **e. Tools:** VS Code, PyCharm, Arduino IDE
- **f. Hosting:** Heroku, Vercel, or AWS

5.3. Functional Requirements

- **a. Collect sensor data** in real time
- **b. Compute RI score** and classify water quality
- **c. Recommend purification methods** based on RI
- **d. Display results** on OLED and web dashboard
- **e. Enable real-time monitoring** and historical tracking

5.4. Non-Functional Requirements

- **a. Fast performance** with results in 2–3 seconds
- **b. Scalable architecture** for future upgrades
- **c. Secure data transmission** and user privacy
- **d. Simple UI** for non-technical users
- **e. Reliable operation** across environments
- **f. Easy maintenance** and firmware updates
- **g. Portable design**

6. Result And Discussion

The analysis confirms that the AI-powered Water Quality Detection and Purification Recommendation System offers significant improvements over traditional water testing methods, particularly in terms of speed, accessibility, and actionable feedback. By integrating real-time sensor inputs with AI-driven analysis, the system computes a Refactor Index (RI) that simplifies complex water quality data into an intuitive classification—Good, Moderate, Poor, or Unsafe. This enables users to instantly understand water safety levels and receive appropriate purification suggestions such as boiling, filtration, or UV treatment. The system’s real-time performance and portability make it especially valuable for households, rural communities, and agricultural settings where conventional lab testing is impractical or delayed.

The use of AI reduces human error and subjectivity in interpreting sensor data, ensuring consistent and reliable evaluations. The web-based interface further enhances usability, allowing users to monitor water quality remotely and track historical trends. However, several challenges remain. The accuracy of recommendations depends on the calibration of sensors and the robustness of the RI algorithm. Environmental factors such as sensor drift, temperature fluctuations, or hardware inconsistencies can affect data reliability. Additionally, while the system

provides purification guidance, it does not yet verify the effectiveness of the applied method, which may vary based on local water conditions.

Another limitation is the lack of large-scale, labeled datasets for training AI models to refine purification logic. Building such datasets requires extensive field testing and community engagement. Furthermore, user trust and adoption may be influenced by the clarity of recommendations and the perceived reliability of the system. Ensuring secure data transmission and protecting user privacy are also critical, especially when integrating cloud-based features.

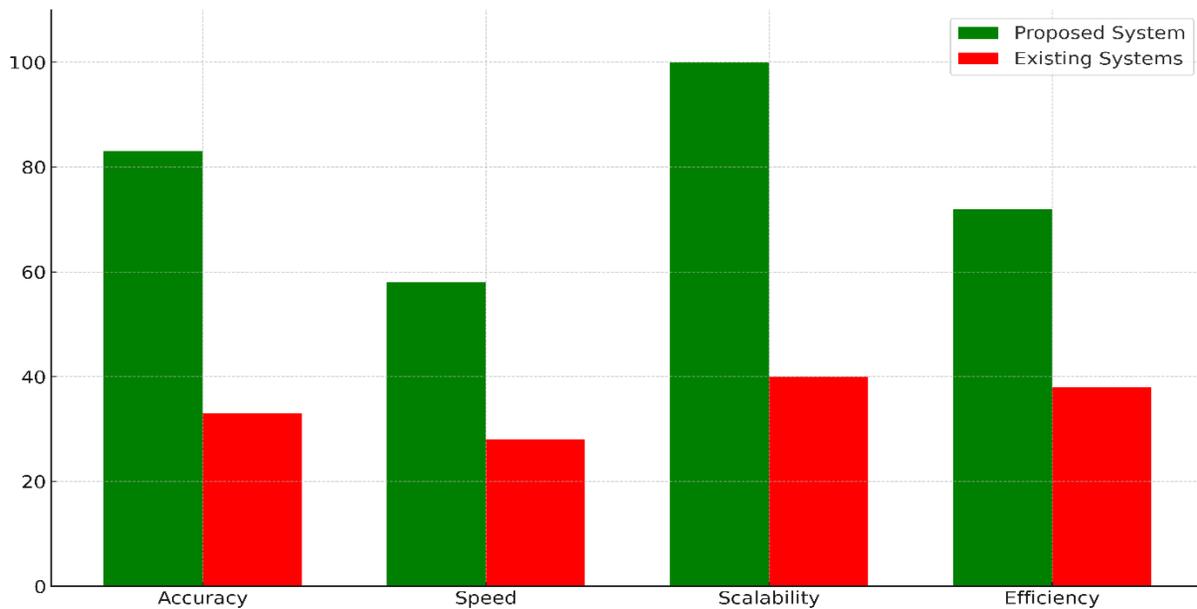


Figure 2: Comparison Matrix

7. References

- [1]. Jin, L., Huang, H., Ren, H. (2025). AI-Driven Transformation of Water Treatment Technology. *Frontiers of Environmental Science and Engineering*.
- [2]. Verma, P., Mehta, P. (2025). Emerging Trends in Real-Time Water Quality Monitoring. *IntechOpen*.
- [3]. Kendir, E., Şerafettin. (2022). Effect of Temperature and Wavelength on Refractive Index of Water. *Indian Journal of Physics*.
- [4]. Qi, Y., He, K. (2025). Science and Technology for Water Purification: Achievements. *MDPI Water*.
- [5]. Ramakant, S., Manvi, S. (2025). Recent Developments in Water Purification. *International Journal of Advanced Chemistry Research*.
- [6]. Ansari, M. et al. (2024). AI for Clean Water: Efficient Water Quality Prediction. *Water Practice & Technology*.
- [7]. Gutenson. (2015). *Water Expert: A Rule-Based Decision Support System for Water Decontamination*. *Drinking Water Engineering and Science Discussions*.
- [8]. Shaghghi. (2024). *DOxy: A Dissolved Oxygen Monitoring System*. *MDPI*.



- [9]. Shravan Kumar, B., Rohith, G., Sai Balaji, A., Tanishq, E. (2025). Water Quality Monitoring System Based on IoT. International Journal of Engineering Technology Research Management (IJETRM).
- [10]. Rosle, N. A. B., Alias, B. (2024). Low-Cost IoT System for Turbidity Measurement. IEEE.