

Integrated Internet of Things and Artificial Intelligence System for Real-Time Multi-Nutrient Water Quality Analysis in Agriculture [†]

Pradeep Kumar Mahapatro , Rasmita Panigrahi *  and Neelamadhab Padhy 

Department of Computer Science and Engineering, School of Engineering and Technology, Gandhi Institute of Engineering and Technology University (GIETU), Gunupur 765022, India; pradeepmahapatro@gmail.com or pradeepkumar.mahapatro@gietu.edu (P.K.M.); dr.neelamadhab@gietu.edu (N.P.)

* Correspondence: rasmita@gietu.edu

[†] Presented at The 11th International Electronic Conference on Sensors and Applications (ECSA-11), 26–28 November 2024; Available online: <https://sciforum.net/event/ecsa-11>.

Abstract: Background: Drinking water that is clean and safe is important for everyone's health. About 1.4 million deaths worldwide occur due to contaminated drinking water each year. Contaminated water sources are the primary cause of diarrheal infections, which account for about 505,000 deaths every year. To overcome these challenges, this work proposes an integrated IoT- and AI-based solution for real-time multi-nutrient water quality analysis. **Objective:** In this paper, we aim to develop a complete integrated with an IoT-based water nutrient analysis system using advanced machine learning models that can predict multiple nutrient levels to improve crops and increase the interpretability, reliability, and security of water quality monitoring systems. **Material/Method:** For data collection, we deployed IoT sensors in different water sources like reservoirs, irrigation canals, and ponds for continuously monitoring parameters including phosphorus (P), potassium (K), pH, temperature, and BOD. The data that we collected from the sensors were securely transmitted to a cloud-based platform using end-to-end encryption protocols. Advanced machine learning classifiers and ensemble learning algorithms were used to analyze the real-time data to produce multi-nutrient predictions. The dataset was collected from GIETU agricultural fields over 6 months from 1 January to June 2024. We also used explainable AI (XAI) techniques to interpret the machine learning algorithms. **Result:** Performance metrics like accuracy, precision, recall, and F1-score were calculated for the water quality prediction. Our experimental observations revealed that the RFS ensemble classifier (Random Forest + SVM) performed well in comparison to other models and had an accuracy of 90%. The hybrid classifier was significantly better than the traditional approaches. We also used XAI techniques to increase the interpretability of the classifiers to enable effective decision-making for water management. For data security, we used encryption and decryption algorithms to ensure data integrity and protection from unauthorized access.

Keywords: soil nutrient analysis; IoT-based water quality monitoring; machine learning; secure data transmission; cloud server



Citation: Mahapatro, P.K.; Panigrahi, R.; Padhy, N. Integrated Internet of Things and Artificial Intelligence System for Real-Time Multi-Nutrient Water Quality Analysis in Agriculture. *Eng. Proc.* **2024**, *82*, 72. <https://doi.org/10.3390/ecsa-11-20358>

Academic Editor: Francisco Falcone

Published: 25 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The demand for food supply is continuously increasing due to the increase in population, but, at the same time, supply is decreasing. So, to increase crop yield, we need to supplement the soil with the appropriate amount of nutrients like nitrogen (N), phosphorus (P), and potassium (K). There are several methods, both physical or chemical, that can be adopted to improve the nutrient level of soil. Additionally, optical methods are suitable for the detection of soil nutrients using sensors. Sustainable farming methods are essential to the global economy, and managing water quality effectively is crucial. Crop health and productivity are directly impacted by proper water quality management, particularly in

contemporary agricultural systems that significantly rely on precise fertilizer management. In the past few years, advances in technology have made it possible to monitor and keep water conditions in farms at their best. Artificial intelligence (AI) and Internet of Things (IoT) devices have revolutionized traditional agriculture by automating procedures, enabling predictive decision making and offering real-time insights into water quality. The purpose of this work is to construct a real-time multi-nutrient water quality measurement system for agriculture that incorporates IoT and AI.

1.1. Background on IoT and AI Systems in Agriculture

IoT systems are widely used in agricultural systems for remotely monitoring and managing farming systems through different sensors. These sensors collect data from different sources and provide real-time feedback to the system. These sensors measure critical parameters such as temperature, humidity, water quality, and soil moisture. Additionally, these data are sent through wireless networks to a hub. Monitoring water quality is especially for making sure that farming methods are appropriate since water and uneven nutrient levels can harm crop growth.

AI, specifically machine learning (ML), is a frequently used tool in precision agriculture. AI can make predictive models that help farmers optimize their water use, nutrient management, and overall farm output by looking at big datasets. IoT and AI can be used together to make smart systems that can monitor, analyze, and react to changes in their surroundings on their own.

To improve the productivity of agriculture, farm soil analysis is important. For this task, some convolutional chemical analysis techniques have been introduced to give some new approaches to measure the characteristics of the soil via the collection of real-time data on soil parameters. Here, we introduced improved sensors integrated with the IoT (Internet of Things). Using machine learning algorithms like decision trees, CNN, and regression helps to predict crops and fertilizers by taking into consideration soil nutrient data like NPK and other parameters like temperature, pH value, and ground cover percentage. Plants that are already mature need a sufficient amount of fertilizer to rapidly grow. For these plants, we used CNN with digital image analysis to monitor and predict the required amount of fertilizer.

Soil sensors and Arduino are often used to determine the nutrient levels in the soil. Crop fertility is based on how much nutrient is supplied to the plant. If a sufficient amount of nutrients is not supplied, then we can verify the levels of the nutrients by using NPK sensors and Arduino, which can then be applied to the soil.

Water covers more than two-thirds of the Earth's surface and is a critical resource for living organisms. However, despite its abundance, the amount of potable water is limited. Moreover, numerous ailments are transmitted through water; hence, the real-time monitoring of water quality (WQ) is essential. Assessing WQ usually entails collecting water samples from various sites at different time intervals and evaluating them in laboratories. However, the manual sampling and laboratory analysis of WQ for any given water body or process can be inefficient, expensive, and time-consuming. As a result, intelligent systems are increasingly being used to monitor WQ, especially when real-time data are needed.

1.2. Motivation and Problem Statement

- (1) Handling real-time sensor data for predictive modeling.
- (2) Developing a hybrid ensemble learning model for predicting the real-time multi-nutrient water quality in agriculture.
- (3) Using explainable AI (XAI) to ensure the interpretability of the machine learning model for predicting real-time multi-nutrient water quality in agriculture.

1.3. Importance of Multi-Nutrient Water Quality Analysis

The real-time monitoring of these nutrient levels is essential for timely intervention, ensuring that corrective measures can be implemented before nutrient deficiencies or

toxicities occur. The proper growth of a plant relies on essential macronutrients such as phosphorus, potassium, and nitrogen. The absence of any one of these nutrients might result in inadequate crop output or soil deterioration. Typical water quality monitoring systems have various disadvantages, including covering a restricted number of parameters and not providing a comprehensive picture of nutrient dispersion.

Novelty: We explored the literature for IoT or AI systems for water quality monitoring. Our approach is innovative in that it combines both technologies for real-time multi-nutrient analysis. Furthermore, reports of applying an ensemble model (RF + SVM) in agricultural water management are scarce, with our method providing significantly higher prediction accuracy and resilience than solo classifiers.

2. Materials and Methods

Clean and safe drinking water is essential for human health. Contaminated water can lead to waterborne diseases such as cholera, dysentery, and giardiasis. Each year, approximately 1.4 million people die due to contaminated drinking water. Diarrheal diseases due to contaminated water sources cause around 505,000 deaths annually [1]. To maximize the crop yield, a plant must receive an adequate amount of nutrients like nitrogen, phosphorus, and potassium. This paper adapted optical and chemical methods to analyze the soil nutrient level in the field. Portable sensors were used, and soil was tested directly for the level of nutrients. But, this method was affected by several environmental factors, which led to inaccuracy in the results provided by the sensors. In this study, Vis-IR spectroscopy was used to detect nitrogen, phosphorous, and potassium nutrient levels, but the results were poor. In the literature, the problem was solved using pretreatment and calibration methods. However, according to our review, the colorimetric method can be used to develop a portable, cost-effective optical sensor for the detection of soil nutrients [2]. Soil fertility plays an important role in the growth of plants, and it determines the quality of the soil. In this paper, Arduino and soil testing sensors were used to determine the content of the nutrients (nitrogen, phosphorous, and potassium) in the soil. If the soil has a low quantity of a specific nutrient, then the sensors can give information on how much of that nutrient should be added so that plants can grow properly and be productive. An NPK sensor was used to detect soil fertility. Due to the scarcity of data, spectral analysis and the classic wet chemistry method did not give sufficient results. In this research, a model was successfully developed to detect the quality of the soil to enable the wise use of fertilizer. By using this model, farmers can easily determine soil content, and select which crop to produce on the quality of the soil [3]. Soil analysis can improve the efficiency of farms and save time and money. To measure the quality of the soil, various techniques have been adopted; the conventional method is chemical analysis. In this study, sensors were integrated with the IoT to monitor and measure soil nutrients in real time and give up-to-date information. These data were collected using machine learning algorithms such as decision trees, random forests, and CNN algorithms, which helped to build a model to accurately predict crop parameters and fertilizer needs. Here, digital image analysis with CNN was used and applied on mature plants to accurately predict the amount of fertilizer required to grow the plants [4].

Traditional farming was performed using the available natural resources like soil, water, and weather. However, it is difficult for the farmer to predict the crop that is suitable for their field's soil conditions. With advancements in the technology in the agriculture sector, it has become easy for farmers to perform tasks from crop selection to harvesting. Machine learning tools, the IoT, and cloud computing help a farmer to analyze data and provide a platform to make better decisions in the process of cultivation. The goal of this research work is to provide a simple and easy way for a farmer to obtain regular information about the field and crop and at the same time to make better decisions at each stage of farming. For this model, AI, ML, cloud, sensors, and automated devices were introduced. In this paper, the IoT-SNA-CR model was introduced to acquire soil nutrient data along with GPS location, moisture, temperature, and water level using sensors [5]. Agriculture is the main

source of India's income. Farmers in India, due to a lack of proper knowledge, make wrong decisions in their farming, which leads to lower productivity. The use of fertilizer plays a key role in agriculture. Farmers think that using more fertilizer increases productivity but plants only use what is required and leave the remaining fertilizer in the soil. The excess application of fertilizer creates many problems regarding soil fertility. So, to avoid this problem, in this paper, we used pre-prepared capsules to test different nutrients like sodium, potassium, and phosphorous [6]. We used a TCS3200 color sensor (manufactured by Texas Instruments, Dallas, TX, USA, 1998), an Arduino microcontroller (developed by Arduino, Turin, Italy, 2005), and soil testing capsules (procured from Elico Limited, Hyderabad, Telangana, India, 2015) to test the soil nutrients and prepared a platform for the farmer to easily make decisions at a lower cost.

For successful farming, there are many methods for estimating soil properties like pH, soil texture, and the C and N present in the soil. Based on the data, one can easily make decisions and select a particular crop to cultivate. This review gives details about the electromagnetic, conductivity-based, and electrochemical techniques for estimating the soil nutrients and pH levels in the soil. Spectral analysis is a method for estimating the availability of nutrients and pH in the soil [7]. In agriculture, soil analysis is key to determining the quality of the soil. So, in precision farming, soil analysis takes place, and a large amount of data must be analyzed to gather information about the quality of the soil. In this study, real-time sensors were deployed in the field and integrated with IoT to continuously monitor and estimate soil nutrients like NPK. Here, MEMS technology was developed to collect data about NPK and other parameters like temperature, pH value, and ground cover percentage. Using machine learning algorithms such as decision tree, CNN, and regression, we analyzed the data collected and prepared a model to predict the suitable crops to grow and fertilizer to increase productivity [8]. Soil analysis gives detailed information about the soil like soil properties and deficiencies. This information is essential to know how to improve the soil quality. If any deficiency is found in the soil, we can supply the required inputs to the soil and increase productivity. Here, we suggested apt treatments to enhance soil fertility, and we collected soil samples from Amity University, Dubai, that we treated with agrochemicals and observed their impact on soil nutrient content and soil pH. Finally, we observed that agrochemical treatment was the best method to improve the soil condition in that region [9]. Soil is the source of nutrients for a plant. We focused on the physical and chemical analysis of the soil and tested the soil to determine the efficiency level of the soil, analyzed the data, and corrected the deficiencies. Soil is influenced by climate, topography (elevation, orientation, and slope of terrain) organisms, and parent materials over time. Improve the quality of soil is a continuous process and achieved through different physical, chemical, and biological processes. This research aimed to provide a platform for a farmer to select a crop, evaluate soil data, and supply what the soil requires to increase productivity [9]. Soil nutrient analysis is an important task to ensure the healthy growth crops. Soil analysis can predict and determine the amounts of nutrients (N, P, K) required for the soil. ML algorithms like classification, regression, and SVM are used to determine the N, P, and K composition in the soil. This method benefits the farmer by ensuring the right crop is grown and the fertilization is appropriate to improve productivity. In this research, MLR (multiple linear regression) models were adopted, and they were 78% accurate in predicting suitable crops with high productivity.

From the above literature survey, we concluded the following:

- Nutrient analyses have been limited scope;
- Predictive modeling is lacking;
- Real-time applications are required.

Proposed Model: Our suggested system integrates IoT sensors with AI algorithms to enable the real-time monitoring and prediction of multiple nutrient levels in agricultural water sources. Our approach is unique in that it incorporates an ensemble model (RF + SVM), which has not previously been used in this domain and provides higher performance in predicting nutritional imbalances.

Phase#1: This is the phase where we collected the data from different sources through sensors for the real-time monitoring of various water quality parameters, including multiple nutrient concentrations (e.g., nitrogen, phosphorus, potassium), pH, electrical conductivity (EC), and temperature. The key components of the system include multi-nutrient Sensors. These sensors monitor the levels of important nutrients (N, P, and K) in the water, which are necessary to sustain ideal farming conditions. pH sensors are used to monitor the water's acidity or alkalinity, which is important for crops to absorb nutrients. EC sensors are used to determine the water's electrical conductivity to gain information about its salinity and general nutrient content. Temperature sensors monitor the water's temperature, which affects the nutrients' solubility and the biological activities of plants. In this phase, we collected data from different sources through sensors like P, K, pH, temp, BOD,. All the data were collected in a real-time environment. These data were fed into the machine learning model as an input, and it is represented in Figure 1. We collected the real-time data from the IoT sensors and stored them on a cloud-based platform. For further analysis. The MQTT (Message Queuing Telemetry Transport) protocol was employed to guarantee dependable data transfer between the cloud infrastructure and the sensors.

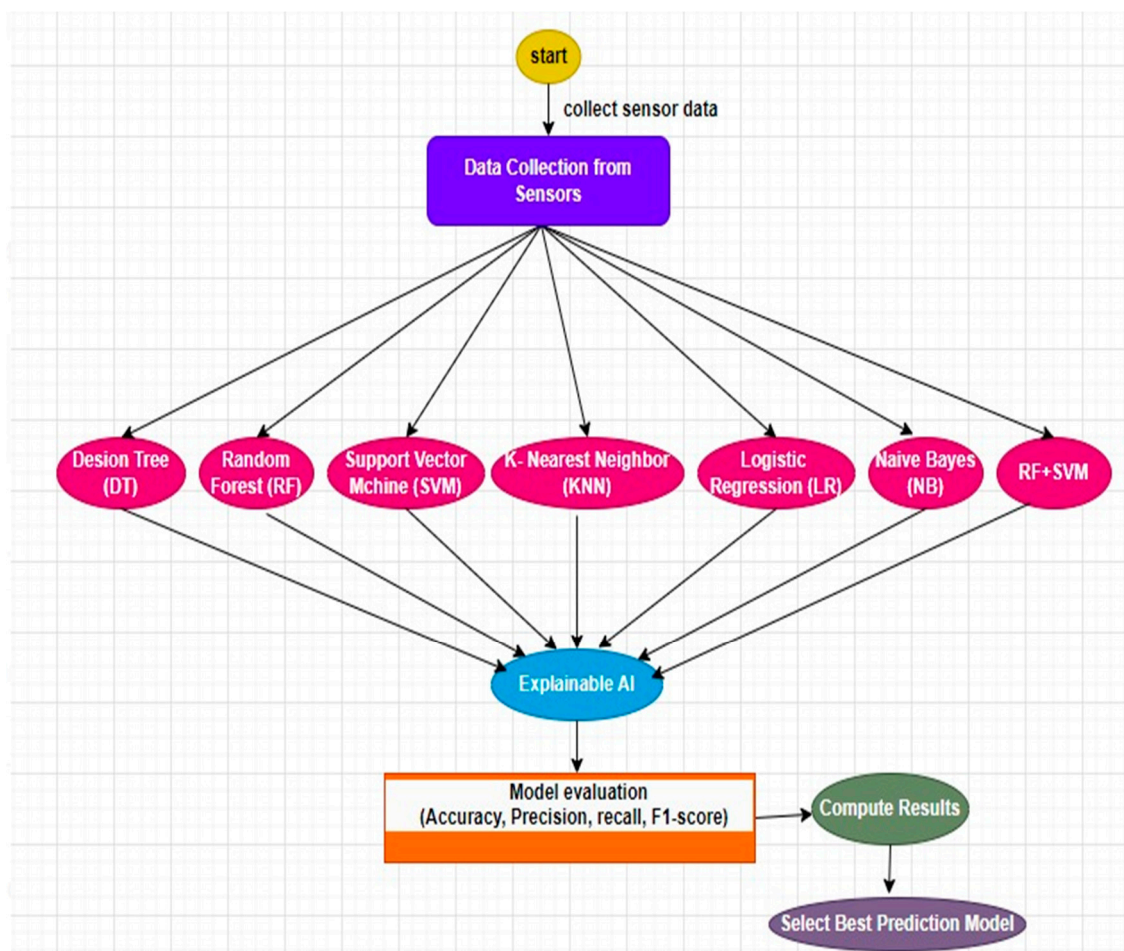


Figure 1. Proposed model for Real-Time Multi-Nutrient Water Quality Analysis in Agriculture.

Phase#2: This phase was the model selection phase, where different machine learning models were identified and trained on the sensor data that we collected in Phase 1. Decision tree (DT) was the classifier used for the sensor data. The model makes a prediction based on the features. Random forest (RF): We used ensemble classifiers that combine multiple DTs that enhance the accuracy and reduce overfitting issues. This model used the collected real-time data and generated the prediction based on the majority voting from the decision tree.

Support vector machine (SVM) gives the best hyperplane that separates both positive and negative classes from the dataset. Similarly, the KNN classifier conducts a classification task by finding the 'K' nearest neighbors and creates the predicted class based on the majority voting. Similarly, the other two classifiers (LR, NB) identify the target classes based on the input features. We used the RFSVM ensemble learning classifiers that combined both classifiers to enhance the accuracy. Phase 3: Explainable AI (XAI): In this phase, we used explainable AI (XAI) techniques for the interpreting and decision-making processes of the models. It explained how the model makes predictions, which helped us to understand the sensor's features. Phase 4: Evaluate models (accuracy, precision, recall, F1-score): In this phase, we estimated the performance metrics of all the classifiers. The metrics (accuracy, precision, recall, F1-score) helped determine which model was the best. Phase 5: Compare Results: In this phase, we compared the performance metrics of the different models and identified the best one. Phase 6: Select the best model: In this phase, we found that our ensemble learning RFSVM model outperformed the other individual models.

3. Result and Discussion

From our experimental observation, we obtain the results presented in Table 1.

Table 1. Experimental observation of different classifiers.

Model Name	Accuracy in %	Precision in %	Recall in %	F1-Score in %
DT	87.30	88.00	87.0 0	87.00
RF	90.40	91.00	90.00	91.00
SVM	55.00	57.00	55.00	50.00
KNN	65.00	67.00	70.00	69.00
LR	70.00	64.00	70.00	65.00
NB	88.00	87.00	88.00	89.00
RF + SVM	90.99	90.0	91.00	92.00

When RF + SVM and individual models were compared, it was clear that the ensemble technique performed better since it incorporated the best features of several classifiers. But, it also added more complexity and required more processing power.

RQ 1: *What makes the performance measures (accuracy, precision, recall, and F1-score) different between classifiers on the same dataset?*

The reason behind this research question was to measure the efficacy of several machine learning algorithms (e.g., DT, RF, SVM, KNN, LR, NB, RF + SVM) when applied to the same dataset for multi-nutrient water quality monitoring. Because of their underlying methods and how they process the data, various classifiers frequently display differences in performance even when utilizing the same data. Performance-influencing factors could include the following:

- (1) **Feature Interactions:** Different algorithms handle feature relationships in different ways. For example, tree-based approaches (DT, RF) inherently capture feature interactions, whereas linear models (LR) do not unless explicitly engineered.
- (2) **Data Distribution:** Some algorithms rely on specific data distributions. For example, NB implies conditional independence, whereas SVM focusses on margin maximisation.
- (3) **Complexity Pattern:** Complex nonlinear patterns can be captured by algorithms such as RF and SVM, although simpler methods such as LR may struggle to do so.
- (4) **Hyperparameter Tuning:** The best hyperparameter settings determine classifier performance. SVM's kernel selection and KNN's k-value can considerably affect results.

To address the above research question, we experimented with and evaluated their performance measures with different measures. The following text provides an overview

of the proposed system integrating IoT sensors with AI algorithms to monitor and predict multiple nutrient levels in agricultural water sources in real-time. The system can enhance farming conditions by analyzing various water quality parameters including nitrogen, phosphorus, potassium, pH, electrical conductivity, and temperature. The collected data are then fed into machine learning models for further analysis.

In phase 1, data were collected from different sources through sensors for the real-time monitoring of water quality parameters. The key components of the system include multi-nutrient sensors, pH sensors, electrical conductivity sensors, and temperature sensors. The collected real-time data were stored in a cloud-based platform for further analysis using the MQTT protocol to ensure reliable data transfer between the cloud infrastructure and the sensors. Phase 2 involved model selection, where different machine learning models such as decision tree (DT), random forest (RF), support vector machine (SVM), KNN, and other classifiers were identified and trained on the sensor data collected in Phase 1. The models were used to make predictions based on the features and enhance the accuracy of the analysis. In Phase 3, explainable AI (XAI) techniques were employed to interpret and understand the decision-making processes of the models. This helped in understanding how the models make predictions based on the sensor's features. Phase 4 involved evaluating the performance metrics of all the classifiers, including accuracy, precision, recall, and F1-score, to determine the best model. In Phase 5, the performance metrics of the different models were compared to identify the best-performing model. In Phase 6, the ensemble learning RFSVM model was identified as the best-performing model based on the experimental observations and performance metrics.

The experimental observation results are presented in Table 1, which shows the accuracy, precision, recall, and F1-score of the different classifiers. The results indicate that the RF + SVM ensemble technique outperformed the individual models in terms of accuracy and other performance metrics, but its complexity is higher, and it requires more processing than the operational classifiers, as shown in Figures 2–4.

The above research question was addressed through the comparison between the precision and recall graphs shown in Figure 5. We selected the variance in performance metrics for all the classification techniques. We also discussed and visualized the factors that impacted the models. The logistic regression and naïve Bayes classifiers are as simple models that could not capture the complex patterns. Other models like SVM and KNN are very sensitive to model building; their performance is affected if they are not scaled properly.

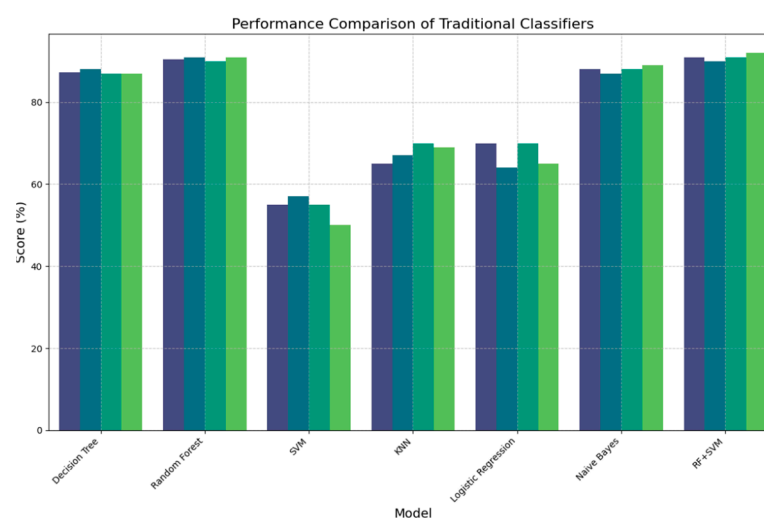


Figure 2. The performance measures of the different classifiers.

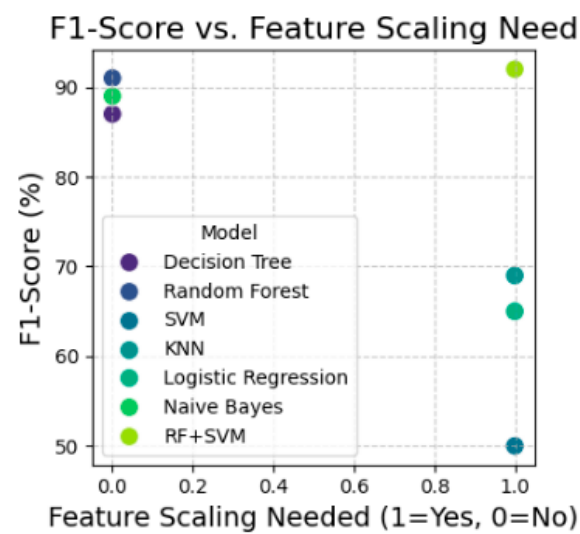


Figure 3. Comparison between F1-Score vs. feature scaling.

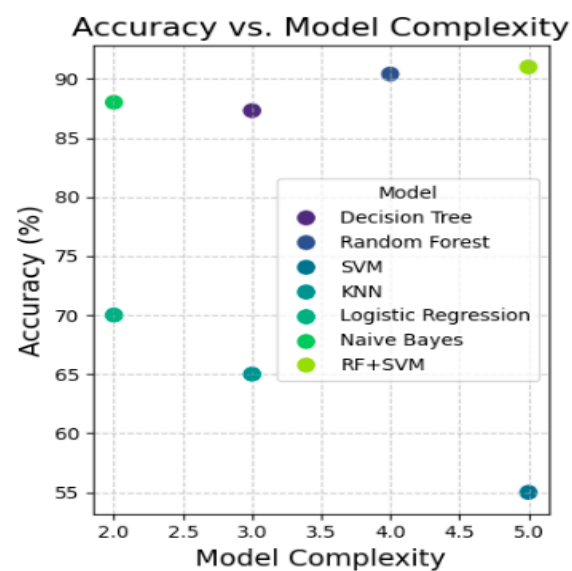


Figure 4. Comparison between accuracy and model complexity.

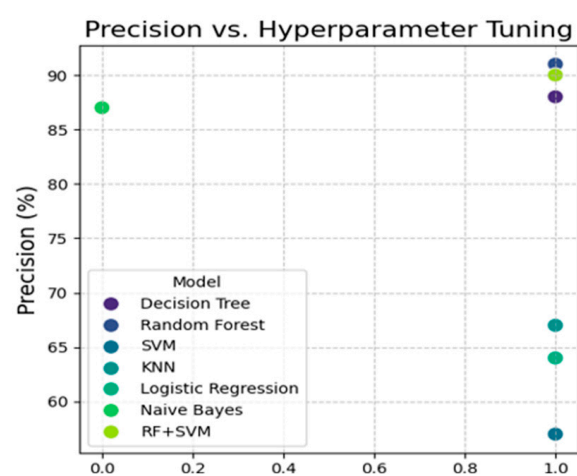


Figure 5. Comparison between precision and hyperparameter tuning.

RQ 2: How does the performance compare of the ensemble model (RF + SVM) with that of the individual models?

Hypotheses:

To address the above research question, we defined the two hypotheses, i.e., H0 and H1

Null Hypothesis (H0): The proposed ensemble model does not perform well in comparison to the individual classifier on the performance metrics.

Alternative Hypothesis (H1): The proposed ensemble model RF + SVM performs well in comparison to others in terms of accuracy, precision, recall, and F1-score.

Our experimental observation was that the proposed ensemble learning (RF + SVM) model obtained higher performance measures than the other traditional classifiers. The obtained accuracy was 90.99%, which can be compared to that of traditional RF i.e., 90.4%. The same was found for the F1-score. The ensemble model's improvement suggests that combining random forest's and SVM's strengths results in stronger predictions.

Figures 6 and 7 compare the accuracy and F1-Score of the classifiers, and Figure 8 presents a boxplot visualization demonstrating the different models' performance metrics. It exhibits the variability in and distribution of these metrics. It shows which classifiers performed constantly well. It was observed that the SVM and KNN classifiers performed well because of their higher variability, and they were considered as more sensitive models for handling real-time data. Finally, it was observed that the RF + SVM model outperformed the other models and was more reliable for handling real-time sensor data.

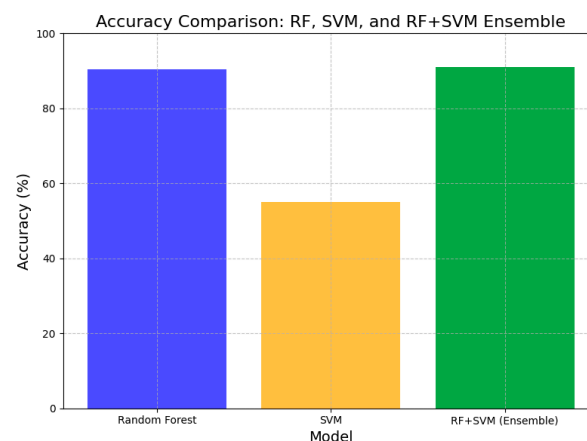


Figure 6. Accuracy comparison of the classifiers.

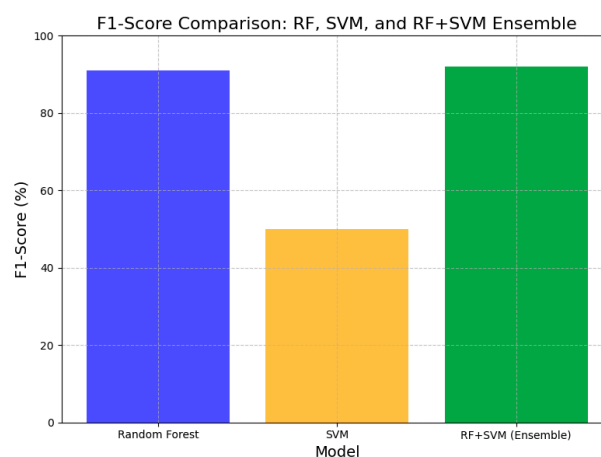


Figure 7. F1-Score of the classifiers.

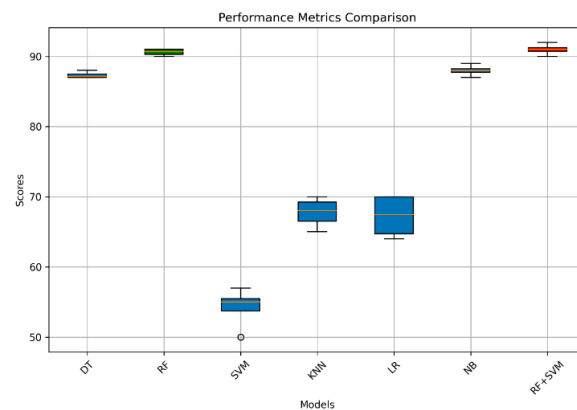


Figure 8. Boxplot representation of the different classifiers.

Figures 9 and 10 demonstrate the trade-off between the precision and recall for each model. From the above observations, we rejected the H0 in favor of H1, suggesting that the ensemble model RF + SVM did indeed significantly outperform the individual models in terms of accuracy, recall, and F1-score. We rejected H0: our proposed ensemble model RFSVM performed well in terms of performance measures (accuracy, recall, and F1-score). We accepted H1: our proposed ensemble model performed slightly better compared to the individual models when considering accuracy, recall, and F1-score.

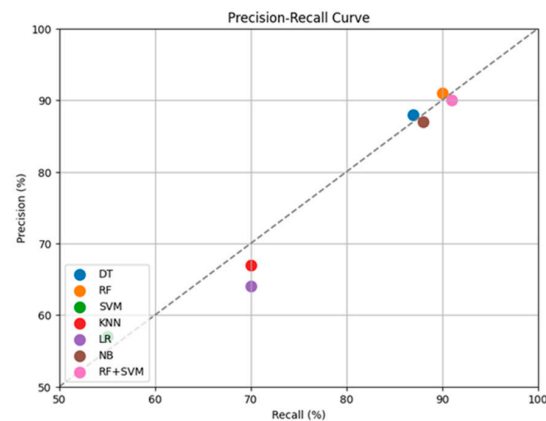


Figure 9. Relationship between precision and recall.

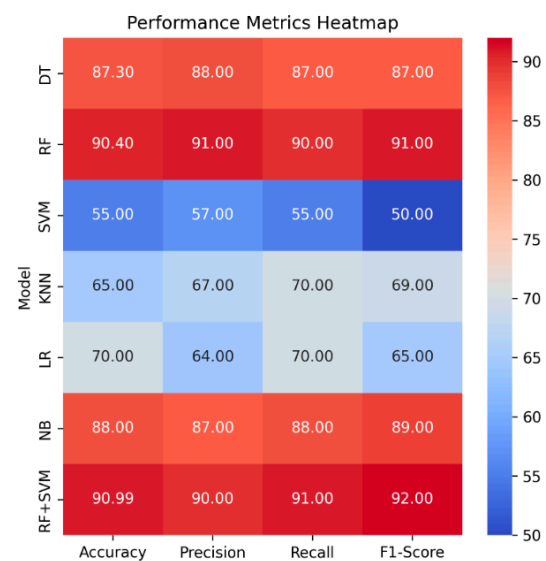


Figure 10. Heatmap representation of all the classifiers.

We used the t -test to determine whether there was a statistically significant difference in performance between the classifiers, specifically between the ensemble model (RF + SVM) and individual models (e.g., decision tree, random forest, SVM). Model performance was evaluated using accuracy, precision, recall, and F1-score, but the t -test let us confirm that the observed differences were statistically significant rather than the result of chance.

Our results were as follows:

H0 (Null Hypothesis): rejected.

H1 (Alternative Hypothesis): accepted.

We used the paired t -test to compare the accuracy of the proposed (RF + SVM) ensemble classifier with that of the individual models.

The test results were as follows:

- The t -statistic values were
 - DT vs. RF + SVM: -76.4194 ;
 - RF vs. RF + SVM: -13.6341 ;
 - SVM vs. RF + SVM: -89.5520 ;
 - KNN vs. RF + SVM: -53.8833 ;
 - LR vs. RF + SVM: $-12,599.0000$;
 - NB vs. RF + SVM: -22.7624 .
- All p -values were 0.0000.

When the p -value < 0.05 , we rejected H0; otherwise, we did not reject H0. Finally, we rejected the null hypothesis (H0) for all comparisons and accept the alternative hypothesis (H1). This means there was a significant difference in accuracy between the RF + SVM ensemble model and each individual model.

Table 2 provides the descriptive statistics. The objective of this table is to present the statistical measures. We estimated the min, max, mean, median, as well as quartiles 1 and 3 for the different metrics for the different classifiers in terms of accuracy, precision, recall, and F1-Score. Figure 11 presents a boxplot of the descriptive statistics; the best results are marked in red.

Our system integrates IoT sensors with AI algorithms to enable the real-time monitoring and prediction of multi-nutrient levels in agricultural water sources. Our approach is unique in that it incorporates an ensemble model (RF + SVM), which has not previously been used in this domain and provides higher performance in predicting nutritional imbalances.

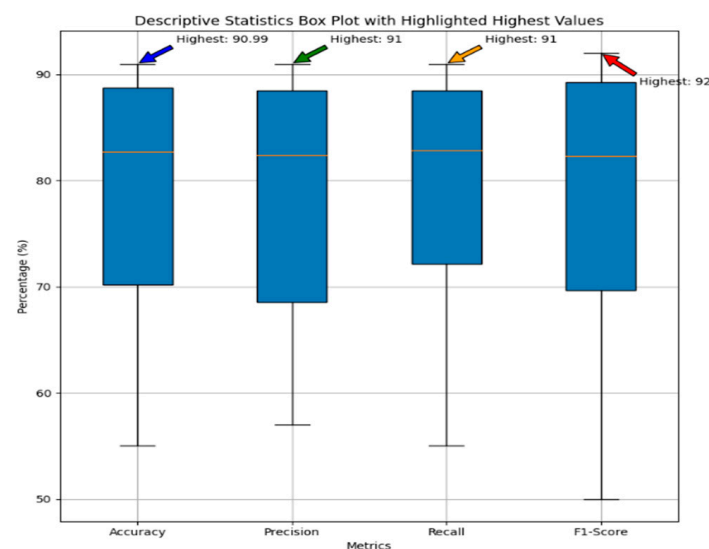


Figure 11. Descriptive statistics boxplot with highlighted best values.

Table 2. Descriptive statistics of all the classifiers.

Statistic	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Min	55	57	55	50
Max	90.99	91	91	92
Mean	78.1	77.71	78.71	77.57
Median	87.3	87	87	87
Q1	67.5	65.5	70	67
Q3	89.2	89	89	90

To determine if integrating models in an ensemble adds value, RQ2 looked into whether this improves the real-time accuracy and predictive ability of the model for multi-nutrient water quality measurement, which is a crucial task in precision agriculture.

Conclusions: In this paper, we divided water quality into four classes based on their “WQI” values: excellent (3), good (2), poor (1), and very poor (0). We used different classification algorithms on this dataset to predict the quality of water. ‘Random Forest’ outperformed the other algorithms in every metric: it achieved an accuracy of 90%, a precision of 91%, a recall of 90%, and an F1-score of 91%. Our study introduced a unique framework for real-time multi-nutrient water quality measurement in agriculture using an integrated IoT and AI system. The use of an ensemble model (RF + SVM) significantly improved prediction accuracy and system robustness, establishing a new standard for nutrient monitoring in precision agriculture.

Author Contributions: Conceptualization, P.K.M. and N.P.; Methodology, R.P.; Software, R.P.; Validation, P.K.M., N.P. and R.P.; Formal Analysis, R.P.; Investigation, N.P.; Resources, P.K.M.; Data Curation, N.P.; Writing—Original Draft Preparation, P.K.M.; Writing—Review and Editing, R.P.; Visualization, P.K.M.; Supervision, N.P. and R.P.; Project Administration, R.P.; Funding Acquisition, N.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used in this study is publicly available and can be accessed from the Kaggle platform.

Acknowledgments: We would like to thank my mentors and the School of Engineering and Technology, Department of Computer Science and Engineering, Gandhi Institute of Engineering and Technology University (GIETU) faculties for their valuable recommendations, professional guidance, expert advice, and encouragement during the preparation of this paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Potdar, R.P.; Shirolkar, M.M.; Verma, A.J.; More, P.S.; Kulkarni, A. Determination of soil nutrients (NPK) using optical methods: A mini-review. *J. Plant Nutr.* **2021**, *44*, 1826–1839. [[CrossRef](#)]
2. Pyngkodi, M.; Thenmozhi, K.; Karthikeyan, M.; Kalpana, T.; Palarimath, S.; Kumar, G.B.A. IoT-Based Soil Nutrients Analysis and Monitoring System for Smart Agriculture. In Proceedings of the 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 17–19 August 2022; IEEE Publications: Piscataway, NJ, USA, 2022; Volume 2022, pp. 489–494. [[CrossRef](#)]
3. Avhad, K.; Mahajan, D.; Kharat, I.; Jadhav, S.; Chattopadhyay, M.M. Soil Nutrients Analysis Techniques and Crop/Fertilizers Prediction—A Review. *Int. J. Eng. Res. Technol.* **2022**, *11*, 12.
4. Senapaty, M.K.; Ray, A.; Padhy, N. IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture. *Computers* **2023**, *12*, 61. [[CrossRef](#)]

5. Pallevada, H.; Parvathi Potu, S.; Venkata Kumar Munnangi, T.; Chandhra Rayapudi, B.; Raghava Gadde, S.; Chinta, M. Real-Time Soil Nutrient Detection and Analysis. In Proceedings of the 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 4–5 March 2021; pp. 1035–1038. [[CrossRef](#)]
6. Najdenko, E.; Lorenz, F.; Dittert, K.; Olf, H.-W. Rapid in-Field Soil Analysis of Plant-Available Nutrients and pH for Precision Agriculture—A Review. *Precision Agric.* **2024**, *25*, 3189–3218.
7. Bhatia, S.; Nair, R.S.; Mishra, V.P. Nutrient Analysis of Soil Samples Treated with Agrochemicals. In Proceedings of the 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, 17–18 March 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 64–67.
8. Wagh, B.; Deshpande, A.; Ingle, R. Study of Physico-Chemical Parameter of Soil Analysis in Buldana District. *Int. J. Adv. Res. Sci. Commun. Technol.* **2021**, *12*, 256–260. [[CrossRef](#)]
9. Madhumathi, R.; Arumuganathan, T.; Iyer, S.; Shruthi, R.; Shruthi, K. Soil Nutrient Analysis Using Machine Learning Techniques. In Proceedings of the National E-Conference on Communication, Computation. Control Autom, (CCCCA-2020), Montreal, QC, Canada, 27 April 2020; Volume 10, pp. 256–260.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.