engineering proceedings

MDPI

Proceeding Paper

Integrated IoT and AI Systems for Real-Time Multi-Nutrient Water Quality Analysis in Agriculture †

Pradeep Kumar Mahapatro [1], Rasmita Panigrahi [1,*] and Neelamadhab Padhy [2]

[1] Gandhi Institute of Engineering and Technology University (GIETU), Gunupur 765022, India; Odisha .School of engineering and technology ,CSE Department
email : pradeepmahapatro@gmail.com
[1*] Gandhi Institute of Engineering and Technology University (GIETU), Gunupur 765022, India; Odisha .School of engineering and technology ,CSE Department
Email : rasmita@giet.edu
[2] Gandhi Institute of Engineering and technology university (GIETU), Gunupur , Odisha .School of engineering and technology ,CSE Department
Email :dr.neelamadhab@hiet.edu

* Correspondence: rasmita@giet.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 are noted to contaminated drinking water each year. Because contaminated water sources are the primary cause of diarrheal infections, they 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, our objective is to develop a complete system that is integrated with an IoT-based water nutrient analysis system using advanced machine learning models that can predict multiple nutrient levels for better crops. To increase the interpretability, reliability, and security of the water quality monitoring system. Material/Method**:** For data collection, we deployed the IoT sensors in different sources like reservoirs, irrigation canals, and ponds for continuously monitoring parameters like:- phosphorus (P), potassium (K), pH, Temperature, BOD, etc. The data that we have collected from the sensors are securely transmitted to a cloud-based platform using end-to-end encryption protocols. Advanced machine learning classifiers and ensemble learning algorithms are used to analyze the real-time data to give multi-nutrient predictions. The dataset was collected from GIETU agricultural fields over 6 months from 2024 January to till date. We also used Explainable AI (XAI) techniques to interpret properly of the machine learning algorithms. Result: The performance metrics like accuracy, precision, recall, and F1-score are calculated for predicting the water quality. Our experimental observation reveals that the ensemble classifier RFS (Random Forest + SVM) classifier exhibits well and has an accuracy of 90% in comparison to other models. The hybrid classifier is significantly higher than the traditional approaches. As well as we used XAI techniques to increase the interpretability of the classifiers to make effective decision-making for water management. For data security, we used encryption and decryption algorithms to ensure data integrity and protection against unauthorized access.

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

Citation: Mohapatra, P.K.; Panigrahi, R.; Padhai, N. Integrated IoT and AI Systems for Real-Time Multi-Nutrient Water Quality Analysis in Agriculture. *Eng. Proc.* **2024**, *6*, x. https://doi.org/10.3390/xxxxx

Academic Editor(s): Name

Published: 26 November 2024

1. Introduction

In the present situation day by day, demand of food supply increases due to the increase in population, but at the same time, supply decreases. So, to increase the yield, we need to feed the soil with the appropriate amount of nutrients like Nitrogen (N), Phosphorus (P), and Potassium (K). There are several methods like physical or chemical adopted to verify the nutrient level of soil. However, the optical method is 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 watch 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 offer a real-time multi-nutrient water quality measurement system for agriculture that incorporates IoT and AI.

Background on IoT and AI Systems in Agriculture

IoT systems are heavily used for agricultural systems for remotely monitoring and managing the farming system through different sensors. The responsible of sensors collect data from the different sources and provide the real-time feedback to the system. The critical measures of those sensors are temperature, humidity, and water quality soil moisture. As well as these data sends through the wireless networks to the Hub. Monitoring water quality is especially important for making sure that farming methods last since water with uneven nutrient levels can have a big effect on food growth.

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

To improve the productivity of agriculture farm soil analysis is the important part. Along with these some convolutional chemical analysis techniques are introduced to give some new approaches to measure characteristics of the soil. Collection of real-time data of soil parameters. Here we introduced improved sensors integrated with 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 etc., For plants that are already grown up they need a sufficient amount of fertilizers to rapid growth. For them, we used CNN with digital image analysis to monitor and predict the required amount of fertilizers.

Soil sensors and Arduino are generally used to determine the nutrient level of 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 level of the nutrient by using NPK sensors and Arduino and can be supplied to the plant.

Water composes more than two-thirds of the earth's surface and is a critical resource for living organisms. However, despite its abundance, the consumable form of water is limited. Moreover, numerous ailments are transmitted through water; hence, real-time monitoring of water quality (WQ) is essential. Commonly, assessing WQ entails collecting water samples from various sites at different time intervals and evaluating them in laboratories. However, 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 used to monitor WQ, especially when real-time data are needed.

Motivation and problem statement:

1. Handling the real-time sensor data for Predictive Modeling
2. A hybrid ensemble learning model developed for predicting Real-Time Multi-Nutrient Water Quality Analysis in Agriculture

3. Explainable AI (XAI) for interpretability of the machine learning model for predicting Real-Time Multi-Nutrient Water Quality Analysis in Agriculture

Importance of Multi-Nutrient Water Quality Analysis

Real-time monitoring of these nutrient levels is essential for timely interventions, ensuring that corrective measures can be taken 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. The typical water quality monitoring system has various disadvantages, including a restricted number of parameters and difficulty in collecting a comprehensive picture of nutrient dispersion. Real-time monitoring of these nutrient levels is critical for prompt interventions, allowing correction actions to be implemented before nutrient deficits or toxicities arise.

Novelty: As per the literature, we explored 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, the application of an ensemble model (RF + SVM) in agricultural water management is unusual, with significantly higher prediction accuracy and resilience than solo classifiers.

2. Materials and Methods

Clean and safe drinking water is essential for every human health. Contaminated water can lead to waterborne diseases such as cholera, dysentery, and giardiasis, from surveying different platforms we have found that due to waterborne. Each year, approximately 1.4 million people die due to Contaminated drinking water. Diarrheal diseases singlehandedly cause around 505,000 deaths annually, due to contaminated water sources [1]. To maximize the production of yield it is necessary that a plant need to get an adequate amount of nutrients like nitrogen, phosphorus, and potassium. This paper adapted optical and chemical methods to analyze the soil nutrient level of the field. Portable sensors are taken and soil is tested directly for the level of nutrients present in the soil. But this method was affected by several environmental factors which leads inaccuracy in the result given by the sensors. In this paper, Vis-IR spectroscopy is used to detect nutrient levels of nitrogen, phosphorous, and potassium but it showed poor results. In the past literature survey in this paper, the problem was solved using pretreatment and calibration methods. However, in the review, it appears that 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 also determines the quality of the soil. In this paper, Arduino and soil testing sensors are used to determine the content of the nutrients (nitrogen, phosphorous, and potassium) in the soil. If the soil is measured with less quantity of a specific nutrient, then the sensors can give information on how much extra nutrient is required to add so that the plant can grow properly and give good productivity. NPK sensor is used to detect soil fertility. Due to the scarcity of data, the result using spectral analysis and the classic wet chemistry method did not give sufficient results. In this research work a model was successfully developed to detect the quality of the soil and use of fertilizer wisely. By using this model an illiterate farmer can easily predict the crop and decide which crop he needs to produce in his field based on the quality of the soil [3]. Soil analysis can improve the efficiency of farms and also save time and money. To measure the quality of the soil various techniques are adopted but the conventional chemical analysis method is one of them. Sensors integrated with IoT to monitor and measure the soil nutrients in real time and give up-to-date information. These data are collected using machine learning algorithms such as decision trees, random forests, and CNN algorithms which will help to build a model to predict accurately the crop and fertilizers. Here in this paper digital image analysis with CNN is used and applied on already grown plants and predict the accurate amount of fertilizer required to grow the plant [4].

Traditional framing was done using the availability of natural resources like soil, water, and weather. However, it is difficult at the farmer's end to predict the suitable crop for his field's soil condition. Now advancement of technology in the agriculture sector it become easy for the farmer from crop selection to crop cutting. Machine learning tools, IoT, and cloud computing help a farmer to analyze the 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 easier way for the farmer who get regular input about the field and crop and also at the same time he can make better decisions at each stage of farming. For this model, AI, ML, cloud, sensors, and automated devices are introduced. Here in this paper, the IoTSNA-CR model was introduced to acquire soil nutrient data along with GPS location, moisture, temperature, and water level using its sensors [5]. Agriculture is the main source of our country's economy. Farmers in our country due to a lack of proper knowledge make wrong decisions in their field which leads to less productivity. The use of fertilizer plays a key role in agriculture. Farmers think that using more fertilizer gives more productivity but plants receive what is required and leave the remaining fertilizer in the soil. Due to excess amount of fertilizer, it creates many problems in soil fertility. So to avoid this problem in this paper researcher used pre-prepared capsules to test different nutrients like sodium, potassium, and phosphorous. Here in this paper researcher used a TCS 3200 color sensor, Arduino, and soil testing capsules to test the soil nutrients and prepared a platform for the farmer to make decisions easily with less cost [6].

For smooth farming, there are many methods for estimating soil properties like pH, soil texture, and C, and N present in the soil. Based on the data one can easily make decisions and predict a particular crop to cultivate. This review gives details about electromagnetic, conductivity-based, and electrochemical techniques for estimating soil nutrients and pH levels in the soil. M. K., [11] conducted the experimental work for real-time data to detect crop disease prediction using machine learning and deep learning classifiers. A complete setup was created to accomplish the task. From a real-time environment, IoT data collected and preprocessed From different classifiers SGD obtained 100% accuracy In agriculture soil analysis is the main part of knowing the quality of the soil. So, in precision farming soil analysis takes place, and a large amount of data to analyze and gather information about the quality of the soil. In this paper, real-time sensors are deployed in the field and integrated with IoT to monitor continuously and estimate soil nutrients like NPK. Here author proposed a MEMS technology 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 they analyzed the data collected and prepared a model to predict a suitable crop and fertilizer to give better productivity [8]. Soil analysis will give detailed information about the soil like soil properties and deficiency in the soil. This information is essential to know how to improve the soil quality if any deficiency is found in the soil so that we can supply the required parameters to the soil and expect better productivity. Here in this paper author suggested apt treatment to enhance soil fertility and the researcher collected the soil samples from form of Amity University, Dubai treated with agrochemicals and observed its impact on soil nutrient content and soil pH. Finally, they observed that agrochemical treatment is the best method to test the soil condition in that region [9]. Soil is the source of supplements to the plant. In this paper researcher focused on physical and chemical analysis of soil and tested the soil to find the efficiency level of the soil, analyzed the data, and corrected the level of deficiency. As soil is influenced by climate, relief (elevation, orientation, and slope of terrain) organisms and parent materials over time. It is a continuous process to improve the quality of soil through different physical, chemical, and biological processes. This research aims to provide a platform for a farmer to predict a crop, evaluate the soil data, and supply the requirements to the soil for good productivity [10]. Soil nutrient analysis is an important criterion for the healthy growth of producing crops. Soil analysis can predict and determine the amount of nutrient composition (N, P, K) required for the soil. ML algorithms like classification, regression, and SVM are used to determine

N, P, and K composition in the soil. This method will benefit the farmer in predicting the right crop and fertilizer for better productivity. In this research, MLR (multiple linear regression) models are adopted, and they give 78% accuracy in predicting suitable crops with a good amount of productivity.

From the above literature survey, we conclude that the following facts

- Nutrient analysis has a limited scope,
- lacks predictive modeling, and
- requires real-time applications.

Proposed Model: Our suggested system integrates IoT sensors with AI algorithms to provide 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 before been used in this domain and provides higher performance in predicting nutritional imbalances.

Phase#1: This is the phase where we have collected the data from the different sources through the sensors for real-time monitoring of various water quality parameters, including multi-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: Keep an eye on the water's acidity or alkalinity, which is important for crops to absorb nutrients. EC sensors: 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 have collected the data from different sources through the sensors like:-P, K, pH, Temp, BOD, etc. We collected the real-time data from the IoT sensors and stored it in the 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

All the data collection is done in a real-time environment. These data are fed into the machine learning model as an input and it is represented in the Figure 1.

Phase#2: This phase is called as model selection phase where different machine learning models are identified and trained on the sensor data that we have collected in Phase 1. Decision Tree (DT): A DT classifier is used for the sensor data. The model makes a prediction based on the features. Random Forest (RF): We have used the ensemble classifiers that combine the 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): This model gives the best hyperplane that separates both positive and negative classes from the dataset. Similarly, the K-NN classifier works for a classification task that finds the 'K' nearest neighbors and makes the predicted class based on the majority voting. Similarly, the other two classifiers (LR, NB) work to identify the target classes based on the input features. We used the RFSVM ensemble learning classifiers that combine both the classifiers to enhance the accuracy. Phase 3: Explainable AI (XAI): In this phase, we used Explainable AI (XAI) techniques for interpreting and decision-making processes of the models. It explains how the model predicts which helps 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 used metrics (accuracy, precision, recall, f1-score) help determine which model is the best. Phase 5: Compare Results: In this section, we compared the performance metrics of the different models and identified the best one. Phase 6: Select Best Model: In this phase, we found that our ensemble learning RFSVM model outperforms in comparison to other individual models.

Result and Discussion: From our experimental observation, we got the result and it is presented in the 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 are compared, it becomes clear that the ensemble technique performs better since it can incorporate the best features of several classifiers. But it also adds more complexity and demands more processing power.

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

The reason behind raising this research question is to measure the efficacy and efficiency 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 is the driving force behind this research question. Because of their underlying methods and how they process the data, various classifiers frequently display varied degrees of performance even when utilizing the same data. Performance-influencing factors could include:

To address the above research question, we have experimented with and evaluated the performance measures with the traditional classifier. The following text provides an overview of a proposed system integrating IoT sensors with AI algorithms to monitor and predict multi-nutrient levels in agricultural water sources in real-time. The system aims to enhance farming conditions by analyzing various water quality parameters including nitrogen, phosphorus, potassium, pH, electrical conductivity, and temperature. The collected data is then fed into machine learning models for further analysis.

In Phase 1, data is collected from different sources through sensors for 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 is stored in a cloud-based platform for further analysis using the MQTT protocol for reliable data transfer between the cloud infrastructure and the sensors. Phase 2 involves model selection, where different machine learning models such as Decision Tree (DT), Random Forest (RF), Support Vector Machine (SVM), K-NN classifier, and other classifiers are identified and trained on the sensor data collected in Phase 1. The models are used to make predictions based on the features and enhance the accuracy of the analysis. In Phase 3, Explainable AI (XAI) techniques are employed to interpret and understand the decision-making processes of the models. This helps in understanding how the models make predictions based on the sensor's features. Phase 4 involves 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 are compared to identify the best-performing model. In Phase 6, the ensemble learning RFSVM model is identified as the best-performing model based on the experimental observation and performance metrics.

The experimental observation results are presented in Table 1, which shows the accuracy, precision, recall, and F1-score of different classifiers. The results indicate that the RF + SVM ensemble technique outperforms individual models in terms of accuracy and other performance metrics, but it also adds complexity and demands more processing power of different classifiers which are represented in Figures 2–4.

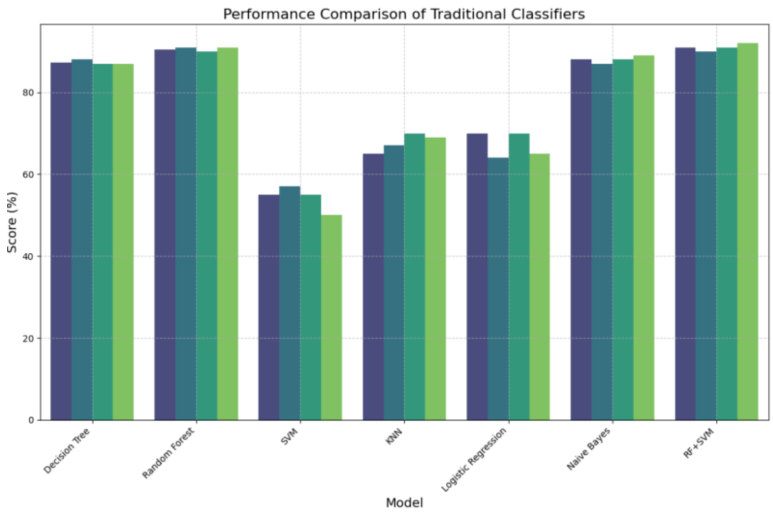

Figure 2. performance measure for different classifiers.

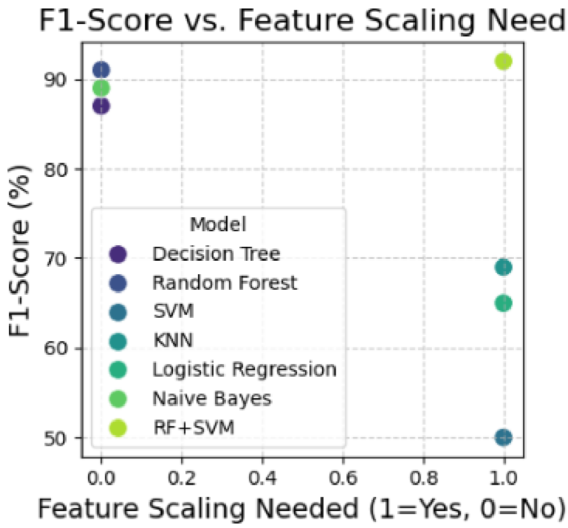

Figure 3. Comparison between F1-score vs. Feature scaling.

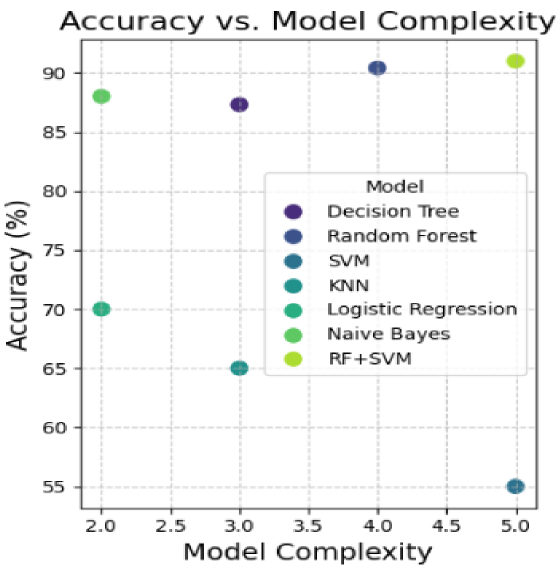

Figure 4. Comparison between Accuracy and model complexity.

The above research question is addressed through the comparison between the precision and recall graph represented in Figure 5. We have implemented the variance in performance metrics concerning all the classification techniques. We also discussed and visualized the factors that impact the models. The classifiers Logistic Regression and Naive Bayes are treated as simple models that suffers to capture the complex patterns. Other models like SVM and KNN are very sensitive towards model building and affect the performance if it is not scaled properly. In Figure.5 presents precision vs. Hyperparameter. Here our objective was to tune the hyper-parameter with 1=Yes and 0=No

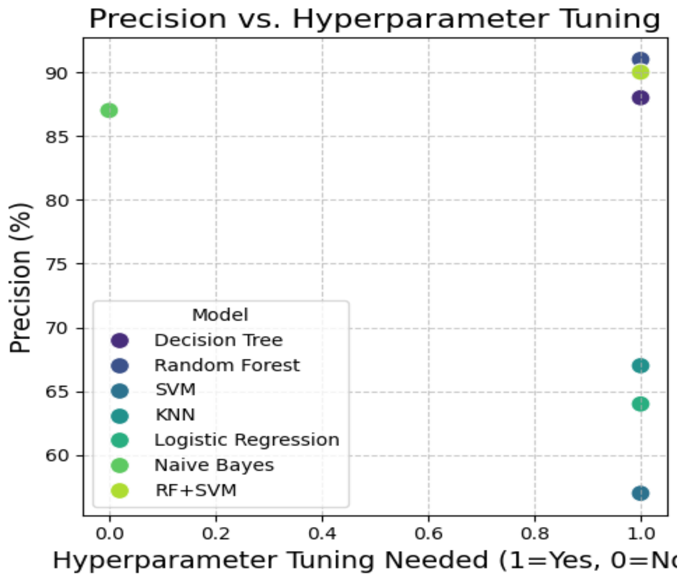

Figure 5. Comparison between precision and Hyperparameter tuning.

RQ 2: *Is it possible to compare the Performance of Ensemble Model (RF + SVM)* vs. *Individual Models?*

Hypotheses:

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

- Null Hypothesis (H0): The proposed ensemble model didn't perform well in comparison to the individual classifier concerning its 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 discussed that the proposed ensemble learning (RF + SVM) model has higher performance measures as compared to other traditional classifiers. The obtained accuracy is 90.99% compared to traditional RF i.e., 90.4%. Similarly, F1-score. The ensemble model's improvement suggests that combining Random Forest's and SVM's strengths results in stronger predictions.

The Figures 6 and 7 represent the accuracy, and F1-score comparison among the classifiers and Figure 8 represents the boxplot visualization which demonstrated the performance metrics across the different models. It exhibited the variability and distribution of these metrics. It discusses which metrics constantly perform well. It has been observed that the classifier SVM and KNN perform well because of higher variability and it is treated as a more sensitive model for handling the real-time data. Finally, it is observed that the model RF + SVM outperforms as a comparison to other models and more reliable for handling the real-time sensor data.

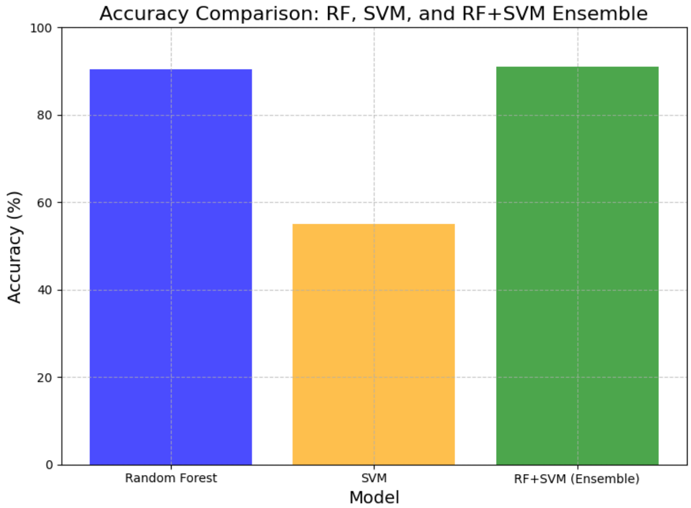

Figure 6. Accuracy comparison among the classifier.

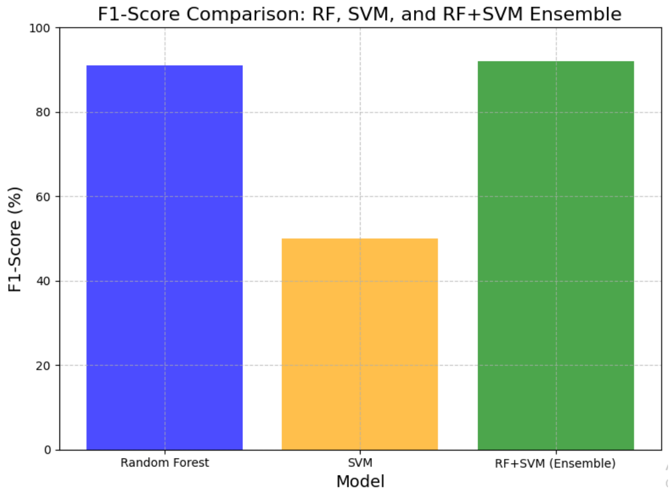

Figure 7. F1-score comparison between all classifiers.

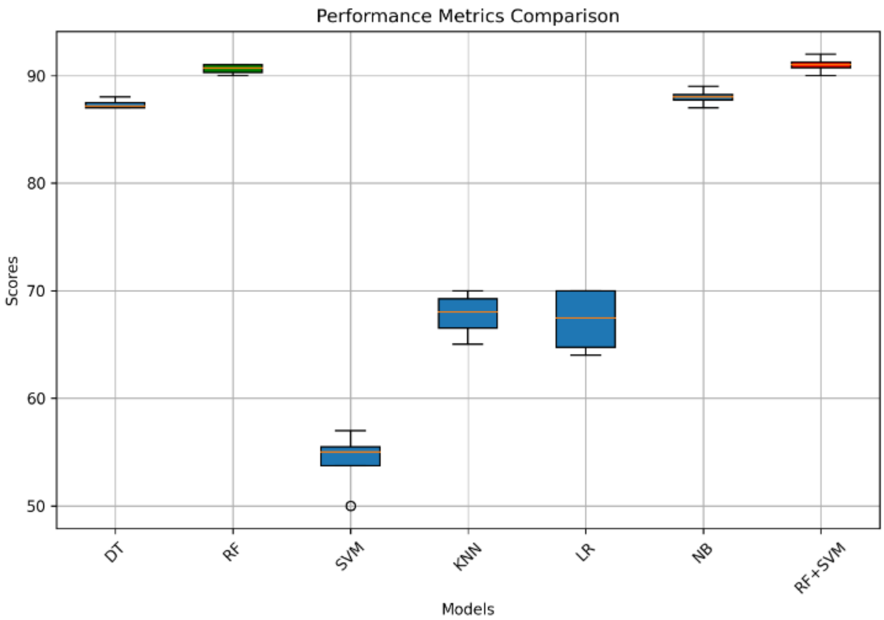

Figure 8. Boxplot representation of different classifiers.

The Figures 9 and 10 demonstrates the trade-off between the precision and recall for each model. From the above observations, it is concluded that we have rejected the **H0 in favor of H1**, suggesting that the ensemble model RF + SVM does indeed significantly outperform the individual models in terms of accuracy, recall, and F1-score. We have rejected the HO: our proposed ensemble model RFSVM performs well in terms of their performance measures (accuracy, recall, and F1-score). Accept H1: Our proposed ensemble model slightly improves the performance compared to individual models when considering accuracy, recall, and F1-score.

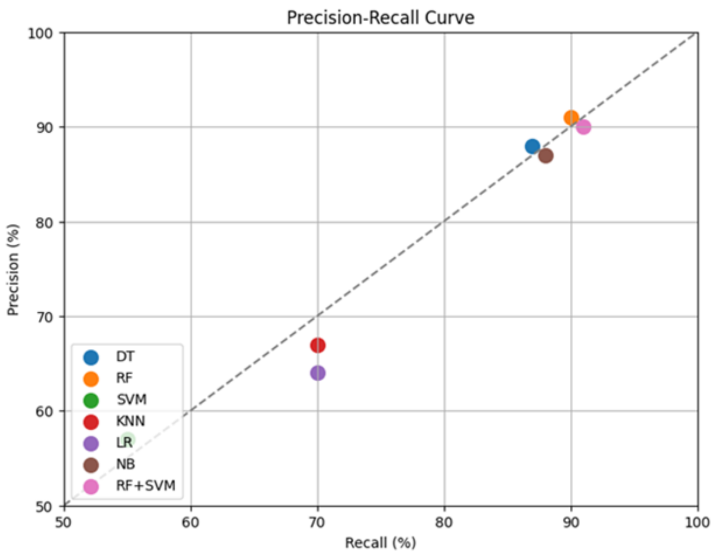

Figure 9. Relation between precision and recall.

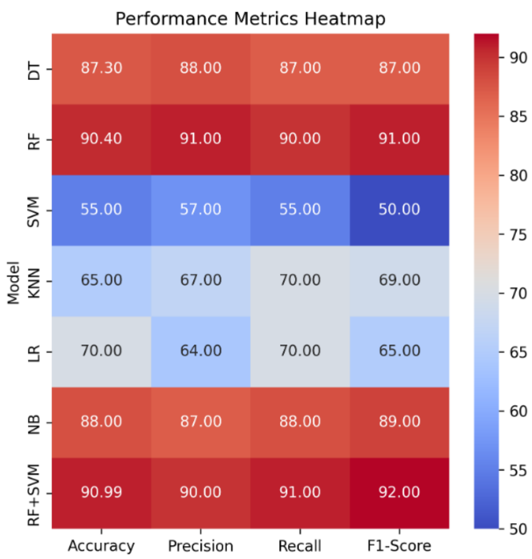

Figure 10. Heatmap representation of all classifier.

The reason behind using the *t*-test is to determine whether there is a statistically significant difference in performance between the classifiers, specifically the ensemble model (RF + SVM) and individual models (e.g., Decision Tree, Random Forest, SVM, etc.), we performed a *t*-test. Model performance is indicated by metrics like accuracy, precision, recall, and F1-score, but the *t*-test lets us confirm that the observed differences are statistically significant rather than the result of chance.

Hence:

H0 (Null Hypothesis): Rejected.

H1 (Alternative Hypothesis): Accepted.

We have used the paired *t*-test to check the accuracy of the proposed (RFSVM) ensemble classifier along with the individual model.

We'll perform a paired *t*-test to compare the accuracy of the RF + SVM ensemble model with each model.

Test Results

- t-Statistic Values:
  - 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
- *p*-Values: All are 0.0000.

when the *p*-value < 0.05: Reject H0, otherwise fail to reject Ho. Finally, we reject the Null hypothesis (HO) for all comparisons and accept the. Alternative Hypothesis (H1). It means there is a significant difference in accuracy between the RF + SVM ensemble model and each model

The Table 2 is the Descriptive Statistics Table. The objective of this table is to present the statistical measure. Here we have estimated the min, max, mean, median, Quartile 1, and 3 for the different metrics for the different classifiers in terms of Accuracy, Precision, Recall, and F1-Score. The below-mentioned Figure 11 represents the descriptive statistics boxplot and is marked as the red color for the highest one.

Table 2. Descriptive Statistics Table of all the classifier.

Descriptive Statistics	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

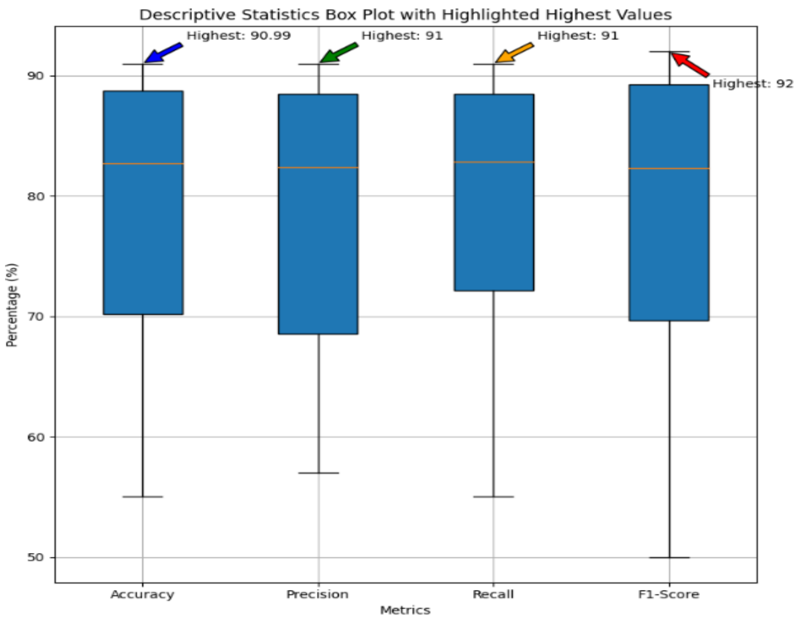

Figure 11. Descriptive Statistics Boxplot representation with highlighted values.

Our suggested system integrates IoT sensors with AI algorithms to provide 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 before been used in this domain and provides higher performance in predicting nutritional imbalances.

In order to determine if integrating models in an ensemble adds value, RQ2 looks into whether doing so improves real-time accuracy and predictive ability for multi-nutrient water quality measurement, which is a crucial precision agricultural task.

Conclusions: In this paper, we divided water quality into different classes based on their "WQI" values where we divided the values obtained into four classes excellent (3), good (2), poor (1), and very poor (0). We used different classification algorithms in this data set for predicting the quality of water from which the 'Random Forest' outclasses every other algorithm in every field, it scored an accuracy of 90%, precision of 91%, recall of 90%, and F1-score of 91%. Our study introduces 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 improves predicted 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 & 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 faculties of department of CSE and School of Engineering and technology 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.
2. Pyingkodi, 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. https://doi.org/10.1109/ICESC54411.2022.9885371.
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. https://doi.org/10.3390/computers12030061References.
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. https://doi.org/10.1109/ICACITE51222.2021.9404549.
6. Najdenko, E.; Lorenz, F.; Dittert, K.; Olfs, H.-W. Rapid in-Field Soil Analysis of Plant-Available Nutrients and pH for Precision Agriculture—A Review. Precision Agric. In Proceedings of the International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), New Delhi, India, 14–15 May 2024; Precision; IEEE Publications: Piscataway, NJ, USA; 2024; Volume 6, pp. 1035–1038. https://doi.org/10.1007/s11119-024-10181-6.
7. Senapaty, M. K., Ray, A., & Padhy, N. (2024). A decision support system for crop recommendation using machine learning classification algorithms. Agriculture, 14(8), 1256.
8. 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.
9. 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.
10. Madhumathi, R.; Arumuganathan, T.; Iyer, S.; Shruthi, R.; Shruthhi, K. Soil Nutrient Analysis Using Machine Learning Techniques. In Proceedings of the National E-Conference on Communication, Computation. Control Autom, (CCCA-2020), Montreal, QC, Canada, 27 April 2020; Volume 10, pp. 256–260.