International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(4s) 2026 pp. 137-157
https://doi.org/10.70917/ijcisim-2026-2499
Article

IoT-Driven Metaverse Agricultural: A Novel Approach for Precision Farming and Sustainability

Phuke Rakesh Rao[1], Ramandeep Sandhu[2], Devanand Shinde[3]

[1]Lovely Professional University, Phagwara, Jalandhar – 144411, Punjab, India.
Email: phuke.rakesh.rao.phd@gmail.com
[2]School of Computer Science and Engineering, Lovely Professional University, Phagwara, Jalandhar – 144411, Punjab, India.
Email: ramandeepsandhu887@gmail.com
[3]Department of Pharmacology, Dr. Babasaheb Ambedkar Marathwada University, Chhatrapati Sambhajinagar (Aurangabad), Maharashtra, India.
Email: devas23@gmail.com
***Corresponding Mail: phuke.rakesh.rao.phd@gmail.com**

Abstract: Due to the current demand for sustainable agriculture and efficient use of resources, more and more modern technologies have been used in agriculture. Issues such as over-watering, over-fertilizing, and sub-optimal yields frequently occur as a result of traditional practices, and are in need of data-driven solutions with exactness. This research introduces an IoT-based metaverse agricultural system for precision farming that leverages real-time monitoring and predictive analytics to enhance the sustainability of agriculture. The sensors are connected to the Arduino Mega microcontroller and the data is transmitted to the cloud through the soil sensors (NPK, pH, moisture, temperature). The data from the sensors is cleaned, smoothed and outliers removed before analysis. For this work, Agriculture IoT 2024 dataset has been selected as it contains different soil types, crop types and environmental parameters which are used for the training and testing of the model. In this paper we propose to use machine learning models, namely MLP, SVM and hybrid ensembles to estimate soil moisture and crop requirements. A hybrid model gives better results with an RMSE value of 0.0229, MAE of 0.0178, and an R-square value of 0.9577. The optimized irrigation and fertilizers led to a 38% reduction in water use and 36% in fertilizer use and to a yield increase of 25%, thus creating a solid, sustainable and effective precision farming methodology.

Keywords: IoT-based Precision Farming, Soil Monitoring Sensors, ML, Sustainable Agriculture

1. Introduction

The field of Agriculture is experiencing a digital revolution driven by the swift advances of cutting-edge technologies such as Internet of Things (IoT), Artificial Intelligence (AI), Robotics, and Immersive Virtual Environments [1]. The need for food source is increasing due to population growth, urbanization, and changing consumption patterns, so that traditional farming alone is not sufficient for the sustainability of agricultural productivity and efficient use of the natural resource [2]. Climate change, soil degradation, limited water resources, increased pest populations and the rising expenses of field operations pose unpredictable problems to modern agriculture [3]. These multifaceted limitations underscore the need to urgently restructure more intelligent, efficient, and sustainable farming systems that are capable of increasing productivity while simultaneously minimizing environmental stress [4]. In this context, the convergence of IoT technologies with the emerging metaverse ecosystem offers a unique opportunity to redefine precision agriculture through real-time monitoring, virtual simulation, and digitally augmented decision support [5]. Figure 1 shows the issues of technology in the agriculture industry.

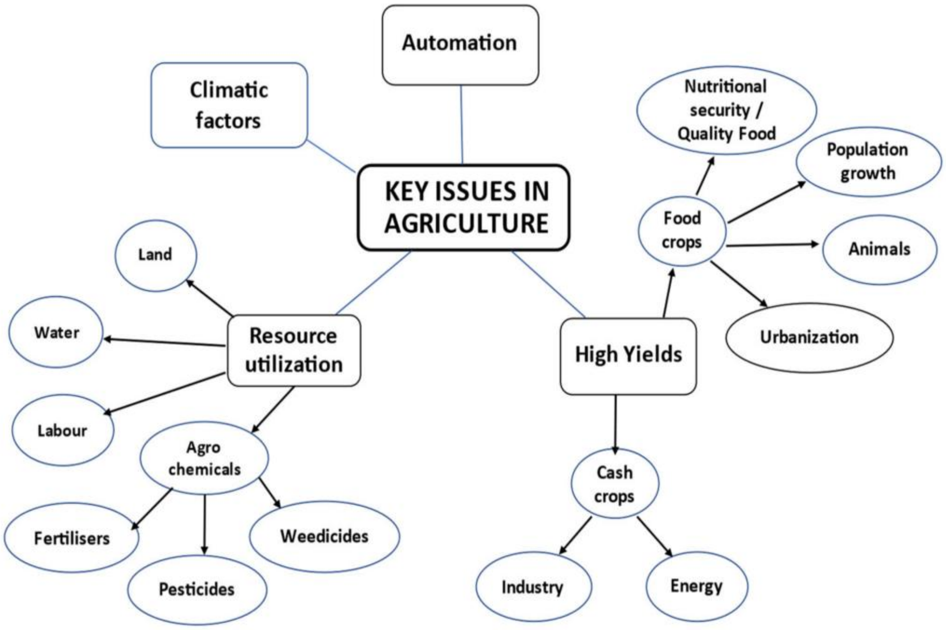


Figure 1: Key issues of technology in the agriculture industry [6].

IoT devices in agriculture have already transformed the functions in the field by facilitating the continuous collection of data through sensors, drones, and automated machines [7]. These sensors produce very fine grained data of such central parameters like soil moisture, soil temperature, nutrient content, plant status, and climate conditions [8]. Nonetheless, these improvements do not always help farmers understand enormous body of data or utilize the insights, because of the complexity of the technical side, a lack of digital fluency, and the inability to create an intuitively designed interface to make decisions in real-time [9]. The metaverse, a networked virtual world that serves as a simulation of real-world entities and processes, is one solution to these constraints because it provides an interactive platform on which data generated by the IoT can be displayed, processed, and responded to. Farmers and stakeholders can gain a clearer insight into crop performance, resource use, and the possible effectiveness of a wide variety of interventions via virtual dashboards, digital twins of farms, and scenario-based simulations [10]. The smart agriculture applications are depicted in figure 2.

Digital twins are a key element in agriculture within the IoT and the Metaverse, where a “digital twin” is a virtual representation of a physical piece of land, its crops, and the surrounding environment using information collected via sensors in real time [11]. This allows users to monitor their crops' health, analyze their soils and evaluate how their irrigation strategies, fertilizer applications or pest control decisions affect their crops without having to disturb their physical fields [12]. By modeling different environmental conditions such as droughts, pests and nutrient needs, agri-businesses have predictive analytics capabilities and can take action proactively at different points during their crop's growth. These capabilities help to lower the risk of crop failure, reduce the waste of inputs, and improve the yields produced, thereby helping to promote sustainable farming practices [13]. When IoT is combined with an Artificial Intelligence (AI) algorithm, they can allow the metaverse environment to be enriched with automated recommendations, anomaly detection and performance forecasting capabilities, thereby creating a data-driven precision farming environment for the future.

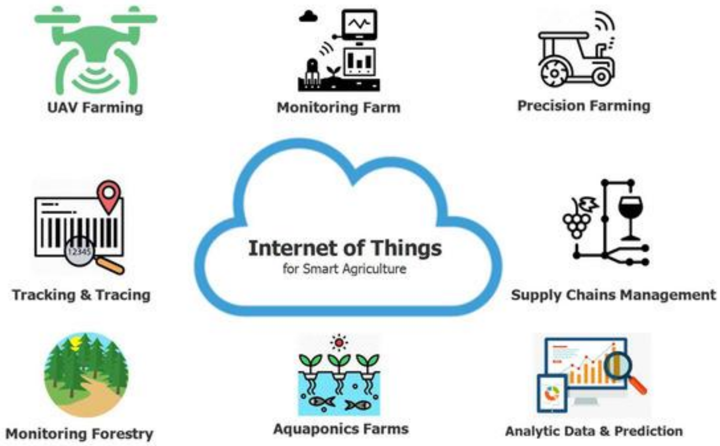


Figure 2: An illustration of IoT applications for smart agriculture [14].

The metaverse can also provide an environment for collaboration and inclusion, allowing farmers, agronomists, researchers, policymakers, and agri-tech companies to easily connect [15]. Knowledge can be shared, issues discussed and new technologies of farming being introduced through virtual meeting rooms, training rooms and interactive learning modules [16] through which agricultural stakeholders can learn. Providing the smallholder farmers with remote advisory services, virtual demonstrations, and on-the-spot consultation and advice from specialists can help them acquire and improve their technical skills without the need to physically travel to other locations [17]. Moreover, financial institutions can leverage the virtual platform for credit and insurance decisions by evaluating the performance of farms, crop health metrics and risk profiles for better financial inclusion of rural communities.

This research explains the possible use of an IoT-based metaverse for precision agriculture through the real-time sensor data and the virtual digital twin of the farm. The system is comprises of soil moisture sensors, NPK sensors, pH sensors and temperature sensors connected to an arduino board which collects field data and pre-process soil data to clean, smooth and remove the outlier. The machine learning models such as MLP, SVM and a hybrid ensemble are created to forecast soil health, irrigation requirement, and fertilizer requirement. The interface of metaverse visualizes the state of the farms, allows making decisions based on simulation, and provides automated recommendations, which can boost the productivity, resource usage, and sustainable management of the farm. Here are the research objectives of the study follows as:

I. To design an IoT-enabled agriculture monitoring system using soil moisture, NPK, moisture, and temperature sensors integrated with an Arduino Mega and cloud connectivity for real-time data acquisition.

II. To develop a metaverse-based digital twin of agricultural fields that visualizes real-time IoT data and enables virtual simulation of crop behavior, environmental interactions, and resource management practices.

III. To preprocess IoT sensor data through cleaning, smoothing, and outlier removal techniques to ensure accuracy, consistency, and reliability of agricultural datasets used for modeling and simulation.

IV. To implement and compare machine learning classification models including MLP, SVM, and a hybrid ensemble model for predicting soil health, irrigation needs, and fertilizer requirements.

V. To evaluate model performance uses metrics such as RMSE, R^2, MAPE, and MAE to determine the most effective algorithm for precision farming applications.

2. Literature Review

Agriculture, being one of the oldest human practices, has undergone a remarkable transformation with the advent of new technologies, especially precision agriculture and precision farming. These newer approaches focus on optimizing crop yields and benefits while minimizing the use of resources. Sharma et al. 2023 [18] highlights technology enabled agriculture to improve productivity by integrating AI and IoT in precision farming. Similarly, Vatin et al. 2024 [19] illustrate the potentials of IoT-driven PA in improving wheat, maize, and soybean yields at 20%, 15%, and 5%, respectively, and reducing pesticide and fertilizer use by 10–20%. Jin et al. 2020 [20], and Panduranga et al. 2024 [21] emphasize the application of deep learning and predictive analytics in order to manage complex data generated in agriculture, predict weather, and automate farm management. Ferrández et al. 2018 [22] and

Bakthavatchalam et al. 2022 [23] further present IoT architecture for soil, weather, and crop condition monitoring along with machine learning techniques to recommend and predictively manage crops with high accuracy in yield prediction and decision support.

Many recent studies have reported on the development of integrated solutions using AI and IoT to produce smart, sustainable, and resource-efficient agricultural systems. Hemal et al. (2025) [24] describe a greenhouse system that uses IoT technology to integrate real-time monitoring, adaptive control and renewable energy sources to increase microclimatic efficiency, determine the optimal types of crops to grow and detect plant diseases with over 97% accuracy and contribute to environmentally friendly practices. Omer et al. (2024) [25] provide evidence of the synergy and effectiveness of using Mobile-Net, support vector machine and K-means clustering to detect crops and weeds accurately with exceptional accuracy and F1-scores. Thilakarathne et al. (2022) [26] and Wongchai et al. (2022) [27] report on the use of cloud-based platform and soft sensors enabled using deep learning to support crop simulation and feature extraction in precision agriculture, providing accurate and timely farm management for multiple crops. In aggregate, these investigations reveal that the incorporation of AI and IoT in precision farming goes beyond just increasing crop yield and saving costs on farm operations to actually enabling eco-friendly, data-driven, and high-tech agricultural practices.

3. Research Methodology

The figure 3 shows an extensive precision farming system powered by the Internet of Things. Various soil sensors are installed at the soil sample level to measure the soil parameters on-line. These sensors have soil moisture, NPK, temperature and pH sensors. All these sensors are connected to an arduino mega microcontroller, to which the sensor data are added and passed to the cloud for the next level instructions. In the middle layer, the data that infers the noise and the smoothing is done for data consistency are subjected to preprocessing after the data cleaning phase. The cleaned data is subsequently introduced to machine learning classification models like MLP, SVM, or hybrid models, which examine patterns and forecast soil health or crop needs. Data is split into 80% training data and 20% testing data to enable evaluation of the model correctly. Finally, the system yields outputs and alerts, providing actionable insights for the farmers to optimize their water, fertilizer and soil usage, thereby boosting yield and resource efficiency.

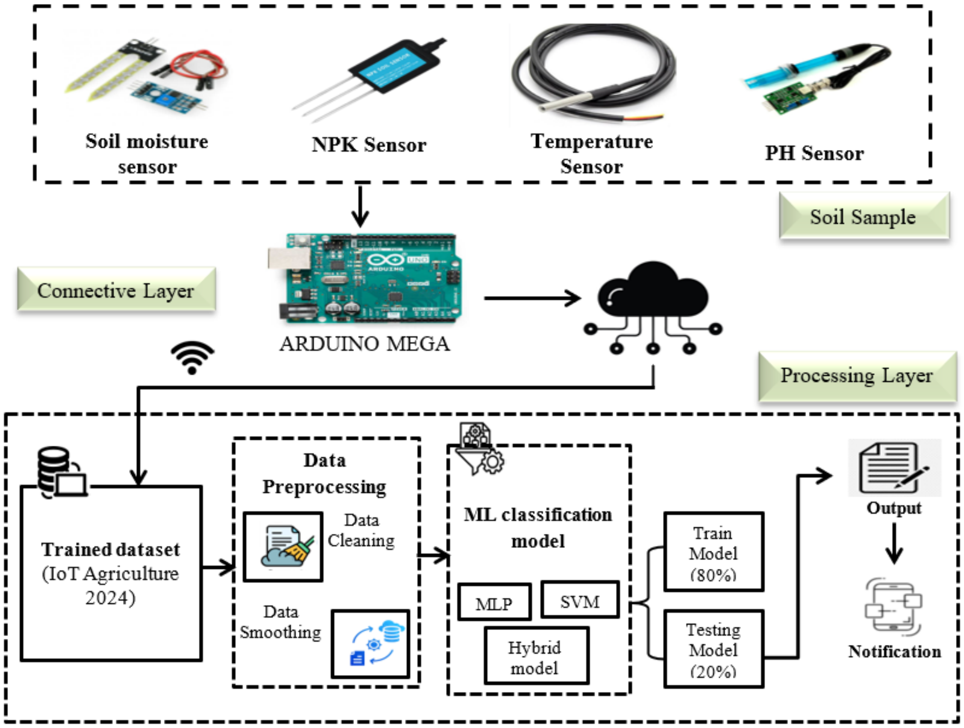


Figure 3: Flowchart of research work

a. Soil Sample

The essential components that form the system suggested for agriculturally monitoring soil probably would be NPK sensor, pH sensor, moisture sensor, and temperature sensor (DS18B20), microcontroller: Arduino Mega 2560, communication module: ESP-01 (Wi-Fi module), cloud storage: for mobile access and uploading of data, and mobile app: end-user interface.

- **NPK sensor**

An electrical conductivity-based sensor measures soil NPK levels by injecting an AC voltage into the soil. The resulting current is variable and reflects conductivity, which depends on the nutrients present. The sensor data are sent to a computer that converts the conductivity values into estimates of nitrogen, phosphorus, and potassium concentrations in the soil. This enables farmers to know exactly how much fertilizer to apply to avoid over- or under-fertilization. It allows farmers to find out where there are not enough nutrients so that they can apply targeted fertilizers in those areas for better crop yields and less environmental impact. Maintaining continuous monitoring will also allow researchers to track how plants absorb nutrients through time, allowing them to make necessary changes to their fertilizing strategies (e.g., type, quantity, frequency) throughout a growing season to ensure that plants grow properly and that their fields are properly managed.

- **Soil moisture sensor**

A moisture sensor is required to measure the water content in soil and is important in characterizing the soil during precision farming. As the water content in soil has a direct influence on plant growth and water access, such sensors can give real-time data to enhance irrigation control. Various kinds of moisture sensors are capacitance sensors which sense the dielectric constant of the soil; tension-meters, which sense the tension of water; and gypsum blocks, which sense the moisture by the electrical conductivity. Through checking the amount of moisture in a field, farmers are able to eliminate over or underwatering, wastage of water and increase crop production. The moisture sensors can also be used to detect the area that is dry or waterlogged so that specific irrigation of the area could be done and overall water management can be done in order to have sustained farming activities.

- **Temperature sensor**

A environment temperature sensor is a device that specifically measures soil temperature, and it is a common tool for soil characterization in precision farming. In general, it employs an analog-to-digital converter to change the analog voltage output to a numeric value that indicates the soil temperature. The digital signal thus obtained can be used for irrigating or for the climate control system to be automated. The accuracy of the sensor is improved by its calibration with a standard or reference temperature. Soil temperature is the main factor that determines seed germination, root development, and nutrient uptake, thus it is the basis of crop performance. The most typical temperature sensors are thermistors, thermocouples, and resistance temperature detectors (RTDs), which identify the temperature-induced changes in the electrical resistance or voltage.

b. Connective Layer

The design of the prototype system utilizes an Arduino Mega with an ATmega2560 microcontroller, which includes four primary sensors: the JXBS 3001 NPK, JXBS 3001 pH, SEN0114 Moisture and DS18B20 Temperature sensors. The NPK and pH Sensors are connected through the RS485 Communication protocol using the MAX485 modules. All four sensors operate on a 12-volt dc supply, and the RS485A and RS485B connections of the sensors are connected to the A and B terminals of their respective MAX485 module. The RE and DE pins of the NPK Sensor are connected to pins 9 and 10 of the Arduino, while the RO and DI pins are connected to the RX3 and TX3. Figure 4 depicts the initial prototype of the system.

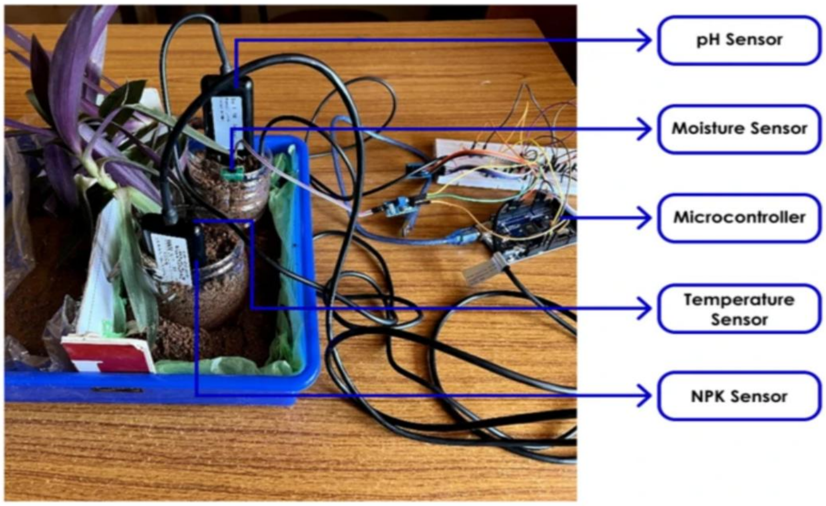


Figure 4: Initial prototype model

The output enables pin (RE) and driver enable pin (DE) of the pH sensor connects to pin 8 and pin 7, respectively, while the receiver output (RO) and digital input (DI) from the pH sensor connect to the data receiving (RX2) and data sending (TX2) pins of the receiving chip. The SEN0114 moisture sensor outputs an analog signal that connects to the A0 (analog input 0) pin of the Arduino microcontroller board, and it is powered via the power (5V) and ground (GND) pins of the Arduino microcontroller board. The DS18B20 digital thermometer uses a 1-wire communication protocol and connects to digital pin six of the Arduino board. The unit contains a 220Ω pull-up resistor on the data line to help stabilize the temperature reading. Finally, the ESP-01 Wi-Fi module uses 3.3V power and connects to RX1 and TX1 of the receiving chip so that you can send data from the sensor to your cloud database in real-time after configuring/running the software through the Arduino IDE.

c. Processing layer

i. Dataset Used

Agriculture IoT in 2024 [28] has certainly progressed towards smarter and more sustainable agriculture, thanks to the combination of connected devices (sensors) that can collect data about the moisture content of the soil, nutrients in the soil, temperature, humidity, and growth of the crops in real time. By providing a constant stream of data, cloud computing and AI allow farmers to make an immediate and accurate decision regarding irrigation, fertilizer application, and pest control. Automated irrigation systems, drone monitoring of crops, and predictive modelling methods for predicting crop yield have all resulted in reduced waste of resources and an increase in crop yield. The incorporation of blockchain technology provides increased transparency and traceability across the complete supply chain. Overall, Agriculture IoT for 2024 has significantly improved efficiency and reduced negative impacts on the environment and contributed towards more sustainable food production methods. Finally, common practice for training/training purposes in a ML application is to use 80% of the training data for training and 20% for testing. Table 1 provides a sample of dataset.

Table 1: Sample of dataset

Temperature	Humidity	Moisture	Soil Type	Crop Type	Nitrogen	Potassium	Phosphorous	Fertilizer Name	Timestamp
26.0	52.0	38.0	Sandy	Maize	37	0	0	Urea	2020-03-06 22:16:11.000000000
29.0	52.0	45.0	Loamy	Sugarcane	12	0	36	DAP	2020-03-06 22:20:14.835729466

34.0	65.0	62.0	Black	Cotton	7	9	30	14-35-14	2020-03-06 22:24:18.671458932
32.0	62.0	34.0	Red	Tobacco	22	0	20	28-28	2020-03-06 22:28:22.507188398
28.0	54.0	46.0	Clayey	Paddy	35	0	0	Urea	2020-03-06 22:32:26.342917864

ii. Data Preprocessing

Preprocessing of the collected IoT sensor data is a crucial step to ensure accuracy, reliability, and smooth integration into the metaverse-based agricultural system. Figure 5 shows the steps of data preprocessing.

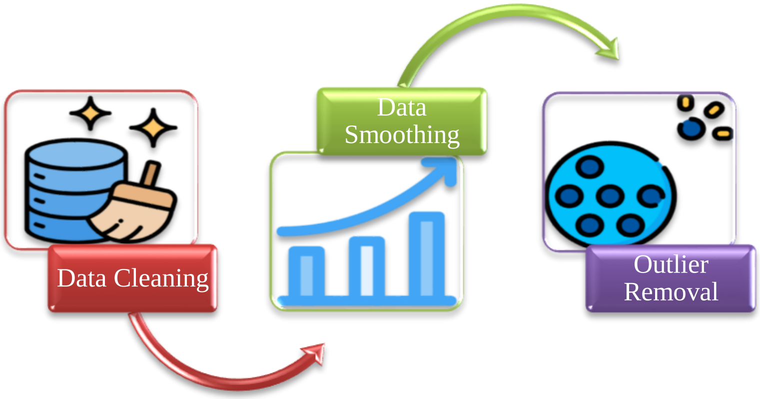


Figure 5: Steps of data preprocessing

- **Data Cleaning**

The first step and the most critical preprocessing step in the collected IoT data is data cleaning that makes sure that the dataset does not contain any errors, inconsistencies, and missing values. The IoT sensors usually give erroneous readings because of connectivity problems, battery depleteness, or other environmental disruptions, and the errors should be rectified prior to any additional analysis. This cleaning operation can eliminate noises generated by missing or invalid values and only meaningful and plausible data remain as a result of this cleaning operation.

- **Data Smoothing**

The dataset is then cleaned and then smoothed to remove random variations or sudden spikes of the data that result from sensor noise, soil effects, or other environmental effects. The process of smoothing enhances smooth transitions and stability of the dataset, which are more appropriate to be used in modeling and visualization within the metaverse. One of the most popular techniques is the moving average technique whereby every value is substituted with the mean of its adjacent readings. This serves to make time series behavior consistent and also gets rid of sudden unrealistic swings, as the data is plotted in the digital twin of the agricultural field.

- **Outlier Removal**

Outlier removal is the final step in the data preprocessing flow, and it mainly focuses on spotting and also correcting those datapoints that are far outside the anticipated agricultural or environmental range. Those outliers can arise from a sensor malfunction, extreme weather irregularities, or just an improper installation, sometimes even small placement issues can do it. By eliminating these outliers, the data becomes more representative of reality, thereby increasing model accuracy and making it less likely to be tricked by AI-based predictions or metaverse-based simulations leading to misleading interpretations.

iii. ML Classification Model

- **Multilayer Perceptron (MLP)**

The multilayer perceptron (MLP) is a supervised learning neural network type that employs the back-propagation approach [29]. Figure 6 demonstrates that three-layer architecture, consisting of an input layer, one or more hidden layers, and an output layer, is advantageous for MLP. In this architecture, every neuron is linked to every neuron in the subsequent layer. It is often reported that MLP possesses significant capabilities in addressing non-linear problems [30,31].

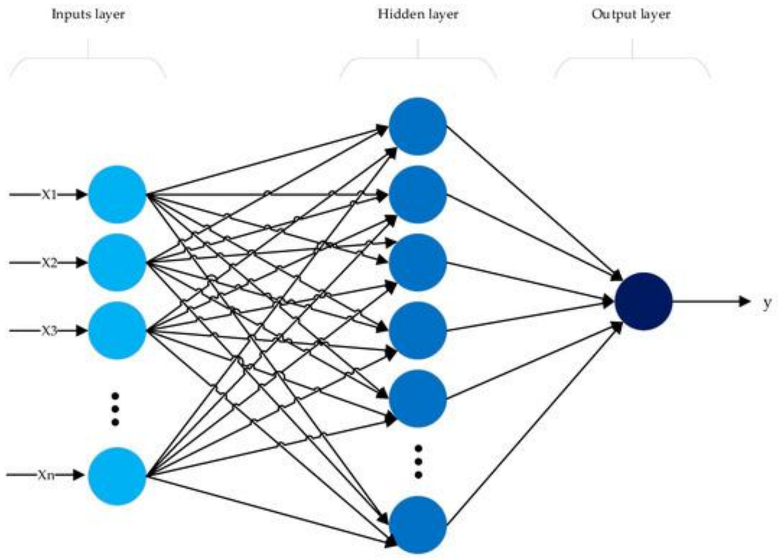


Figure 6: The architecture of the MLP neural networks [32].

The calculation of the output of the input variables, bias values, and input values is shown in Equation (1):

$$S_i = \sum_{i=1}^{n} w_{ij} I_i + \beta_j$$

The input layer is represented by I, the input variable i is denoted as I_i, the total number of inputs is shown by n, β_j is a bias value, and w_{ij} is the weight of connections in the j level [33]. Most multi-layer perceptron activation functions are sigmoid curves, which could be determined using Equation (2):

$$f_j = \frac{1}{1 + e^{-s_j}}$$

in which the activation function is denoted by S. Based on this, researchers could apply Equation (3) to get neuron j final output:

$$y_i = f_i\left(\sum_{i=1}^{n} w_{ij} I_i + \beta_i\right)$$

In order to determine the model's efficacy, it is necessary to compare the MLP method's output value (y) with the target values. As a training dataset, MLP used 70% of the entire data and was randomly sorted by the model [34].

- **Support Vector Machine**

It is one of the powerful supervised machine learning algorithms that are commonly used for classification and regression because of its high accuracy and strong generalization ability [35]. An SVM finds an optimal hyperplane that classifies the data into different classes with maximum margin, which gives better decision boundaries and reduces classification errors [36]. SVM makes use of support vectors, which are important data points nearest to the boundary for defining a hyperplane. This helps in making the model efficient as well as less influenced by noise.

The kernel functions in SVM include linear, polynomial, and radial basis function (RBF) kernels. The input features can be transformed into higher-dimensional feature spaces enabling linear and nonlinear data handling [37]. Due to this property, complex patterns can be captured very effectively. SVM has found its application in precision agriculture, IoT data analysis, crop diseases identification, image classification, biomedical diagnostics, etc., giving reliable performance even with small datasets [38]. Figure 7 shows the classification of SVM.

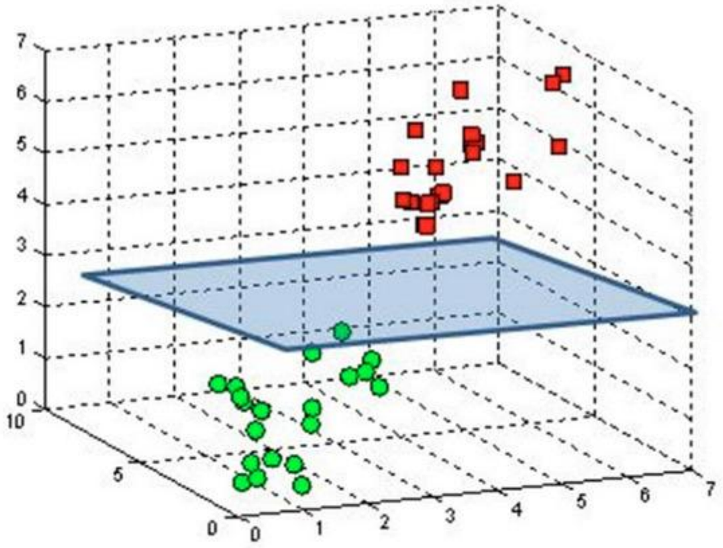


Figure 7: Classification of SVM with Hyperplanes [39].

- **Hybrid Model**

The ML model's ensemble is a common method for improving performance by merging many classifiers. Dataset accuracy could be enhanced by the use of ensemble techniques, which integrate the results of many classifier models. The ensemble frameworks use the two ML models—SVM and MLP—that were already mentioned because they improve the performance of ML-based classifiers and get better results in agriculture.

They calculate each models output, $Y_j, (j = 1,2,3, \dots., m = 6) \in \mathbb{R}^C$ considering C=2 and confidence value $P_i \in \mathbb{R}(i = 1,2)$ on the unrevealed test data where $P_i \in [0,1]$ and $\sum_{i=1}^{C} P_i = 1$. This study presents an approach that utilizes equations to accomplish weighted aggregate of several ML algorithms.

$$P_i^{en} = \frac{\sum_{j=1}^{m=6}(W_i \times P_{ij})}{\sum_{i=1}^{C=2}\sum_{j=1}^{m=6}(W_i \times P_{ij})} \tag{12}$$

The weight of the associated j^{th} classifiers' AUC is denoted as W_j. The output of the ensemble model, $Y \in \mathbb{R}^C$ includes the confidence values $P_i^{en} \in [0,1]$. If $P_i^{en} = \max(Y(X))$, then C_i would be the final class label for the suggested datasets unobserved test data, $X \in \mathbb{R}$, as determined by the ensemble framework.

iv. Performance Metrics

The evaluation parameters for the prediction model are as follows:

$$Water\ Savings\ (\%) = \frac{Water\ used\ (traditional) - Water\ used\ (optimized)}{Water\ used\ (traditional)} \times 100$$

$$Fertilizer\ Reduction\ (\%) = \frac{Fertlizer\ applied\ (traditional) - Fertlizer\ applied\ (optimized)}{Fertlizer\ applied\ (traditional)} \times 100$$

$$Yield\ increases\ (\%) = \frac{Yield\ (optimized) - Yield\ (traditional)}{Yield\ (traditional)} \times 100$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|(x_i - y_i)|$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - y_i}{y_i}\right| \times 100\%$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}$$

4. Result ad Analysis

Utilizing Python, the data on soil moisture, temperature, and nutrients was processed to remove outliers, perform cleaning and smoothing operations. The predictive analysis used SVM, MLP, and Hybrid modeling while Matplotlib and Seaborn provided visualisation and correlation analysis, as well as plotting. Performance metrics were calculated, and custom scripts were created to help with precision farming decision making.

Table 2 provides a comprehensive perspective of the environmental and soil environment that is necessary to facilitate accurate and data-driven agriculture. Tempratures of 26℃ to 34℃ and humidity of 52 to 655 are used to determine the climatic suitability of crops. The soil moisture level is in between 34-62 and it is aided with the help of the moisture sensor readings capturing the changes in the moisture level in order to make correct irrigation decisions. Various types of soils like Sandy, Loamy, black, red, and clayey soils coupled with different crops like Maize, Sugarcane, Cotton, Tobacco, and Paddy bring out different agricultural conditions. Nitrogen, potassium, and phosphorous measurements together with such fertilizers as Urea, DAP, and 14-35-14 provide the information about fertility of soils. In general, such multi-parameter data can be used to support smart monitoring and sustainable precision-farming applications.

Table 2: Load and prepare moisture.

Temperature	**Humidity**	**Moisture**	**Soil Type**	**Crop Type**	**Nitrogen**	**Potassium**	**Phosphorous**	**Fertilizer Name**	**moisture_avg**	**moisture0**	**moisture1**	**moisture2**	**moisture3**	**moisture4**
26.0	52.0	38.0	Sandy	Maize	37	0	0	Urea	0.268	0.33	0.40	0.36	0.23	0.02
29.0	52.0	45.0	Loamy	Sugarcane	12	0	36	DAP	0.482	0.69	0.62	0.69	0.39	0.02
34.0	65.0	62.0	Black	Cotton	7	9	30	14-35-14	0.480	0.68	0.62	0.68	0.39	0.03
32.0	62.0	34.0	Red	Tobacco	22	0	20	28-28	0.478	0.68	0.62	0.68	0.39	0.02
28.0	54.0	46.0	Clayey	Paddy	35	0	0	Urea	0.480	0.68	0.62	0.68	0.40	0.02

a. Environmental Parameters Analysis

Figure 8 shows the distribution graphs of the main environmental and nutrient parameters that were measured in the different locations. Temperature span from 20°C to 40°C, but most of the measurements were around 28–32°C, which means that the climatic conditions were quite stable. Humidity fluctuated from 40% to 80%, thus it formed a very strong central peak near 58–62%, meaning that the atmosphere was moderately saturated with water vapor. Soil moisture varied from 20 to 70, and it had a major concentration around 38–50, which indicates that the soil water content was balanced.

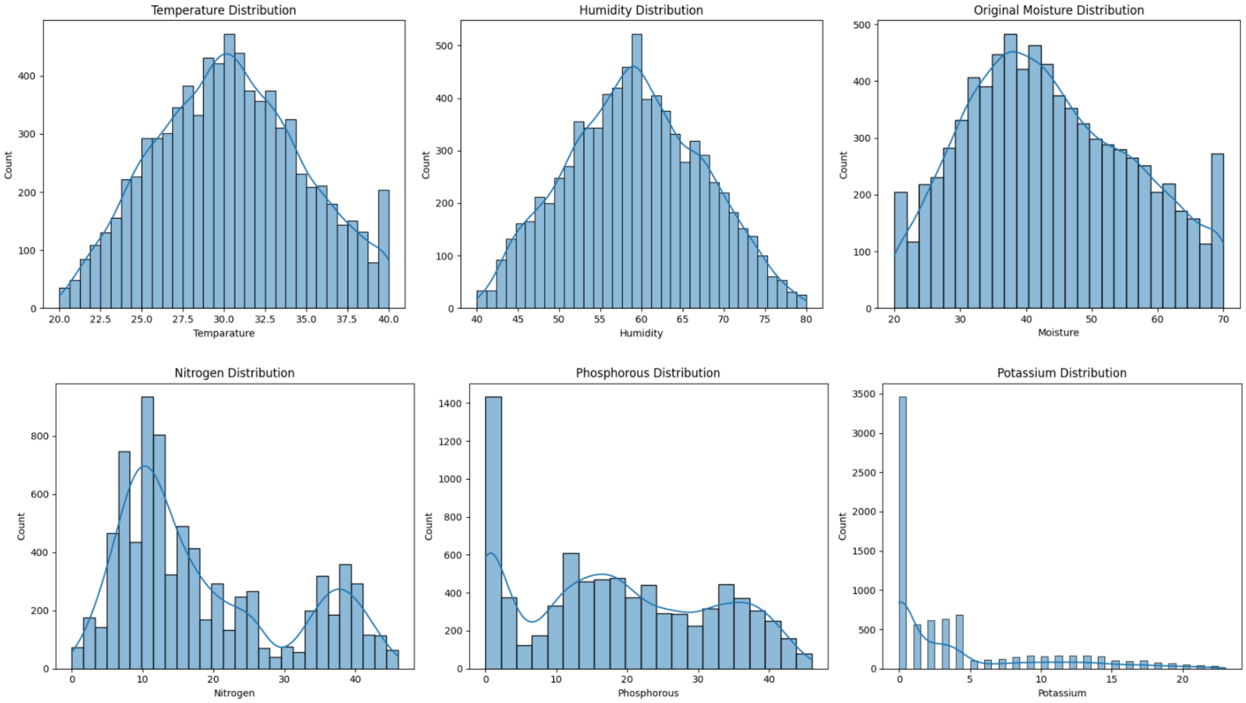


Figure 8: Environmental Parameters Analysis

Nitrogen distribution has several maxima between 5 and 15 with a second increase around 35–45, which can be interpreted as the heterogeneity of the soil fertility. Phosphorous varies between 0 and 45, with substantial accumulations at 5, 20, and 35. Most of the potassium values are quite low, i.e., the bulk is within the range of 0 to 5, while a minor fraction is going up to 22. Such distributions emphasize the diversity of environmental and nutrient situations that have been recorded by IoT sensors.

b. Soil and Crop type distribution

The distribution graph of the soil type exhibits a balanced representation of all the five soil types, that is, Clayey, Black, Red, Loamy, and Sandy, with each type represented about 1550-1650 times (as seen in Figure 9). This even distribution leaves the dataset with varied soil conditions to make sure that the data is modeled and analyzed effectively. Equally, the graph of distribution of crop types has shown the even distribution of all printout crops, such as Sugarcane, Maize, Wheat, Ground Nuts, Pulses, Cotton, Millets, Tobacco, Oil Seeds, Paddy, and Barley with an occurrence of 700-760 times in each category.

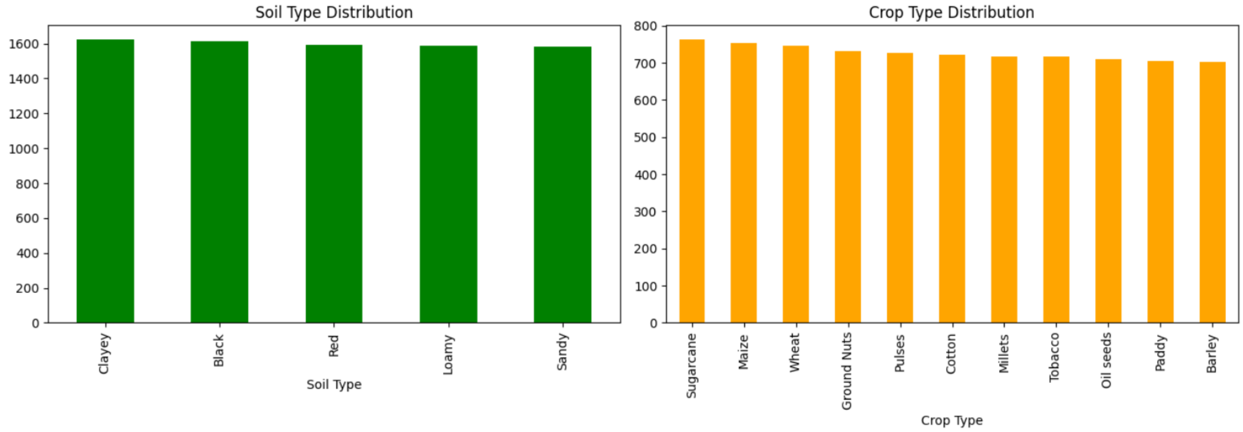


Figure 9: Soil and crop type distribution

This coherent representation shows that the dataset is able to capture a large portion of the agricultural scenario and in the process of generating reliable prediction and generalization across dissimilar crop-soil interactions in precision farming.

c. *Correlation Heat-map*

Figure 10 shows the correlation heat-map relationships between environmental, nutrient, and moisture-related parameters. There is a moderate negative correlation of –0.52 between temperature and humidity; this means that generally, higher temperatures come with lower levels of humidity. Moisture has a weak overall correlation with nitrogen (0.08) and potassium (–0.32), which indicates that soil moisture is more affected by environmental factors rather than nutrient levels.

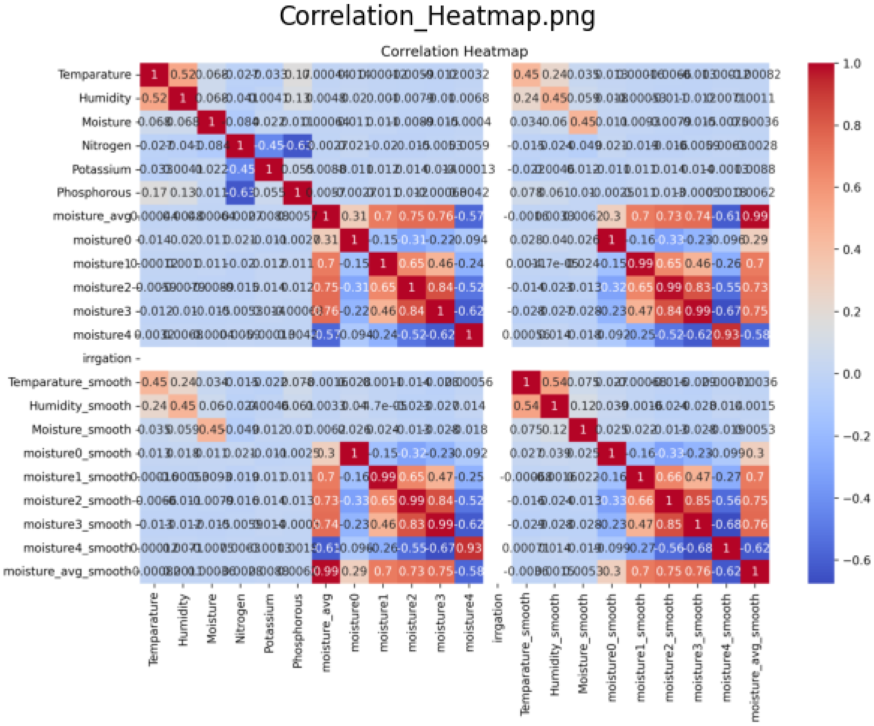


Figure 10: Correlation Heat-map

Almost all the moisture sensors show a strong positive correlation with each other, especially moisture0–moisture1 (0.75) and moisture2–moisture3 (0.84), thus, they confirm that the sensors behaved consistently. The filtered variables are also highly correlated with the original variables as well, the correlation between moisture_avg and moisture_avg_smooth being 0.99, thus, the smoothing performance can be considered stable. In general, the heatmap is revealing significant interactions that are supporting precision agriculture based on data.

d. 3D PLOT of NPK, Temperature, and Moisture

The visualizations of soil nutrients and environmental parameters provide a thorough understanding of the agricultural conditions (as seen in Figure 11-13). Figure 11 shows an NPK pair-plot, which reveals the relationships between Nitrogen, Phosphorus, and Potassium, showing both separate distributions and interaction patterns. Nitrogen is multi-peaked, Phosphorus is more uniformly distributed, and Potassium peaks at lower values, while the off-diagonal scatter plots highlight the nutrient combinations that are grouped in different fields.

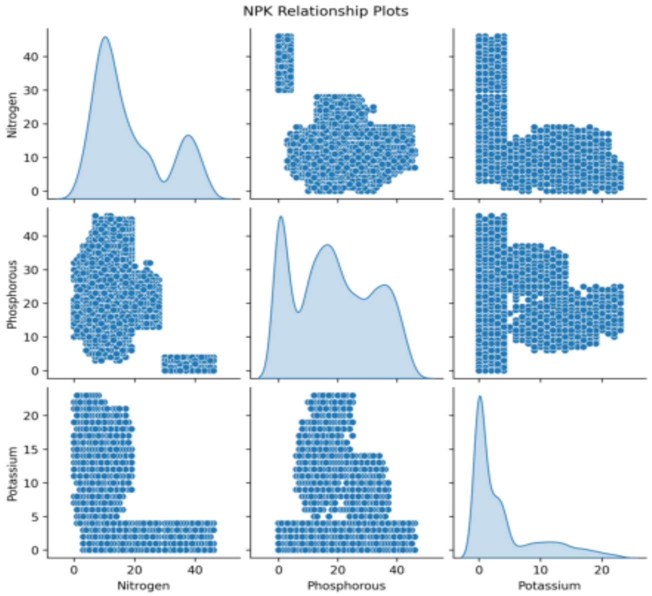


Figure 11: NPK relationship plots

A temperature vs. moisture scatter plot (as seen in Figure 12) that shows how averages of soil moisture (0.2-0.7) are independent of temperatures (20 - 40°C) with some average soil moisture bands occurring more frequently than others and wider soils with higher temperatures show a greater and more diverse distribution of moisture content.

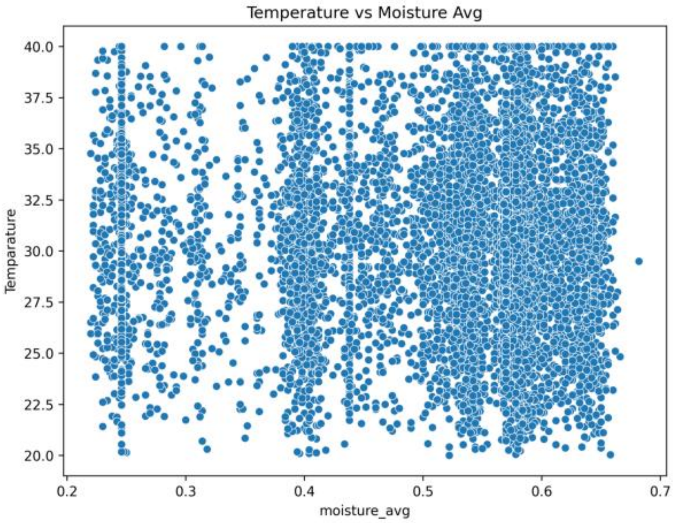


Figure 12: Temperature vs Moisture Avg

In 3D, the humidity-temperature-moisture graph represents multivariate complexity and illustrates dense areas where moderate temperature, higher moisture, and elevated humidity co-occur, while temperature is still widely distributed. Combined, these plots allow a simultaneous examination of nutrient variability in relation to environmental factors which helps to identify clusters as well as interdependent and independent patterns. Such integrated visualization methods can facilitate enhanced understanding of soil conditions for on-site monitoring as well as predictive modeling aimed at efficient agricultural management.

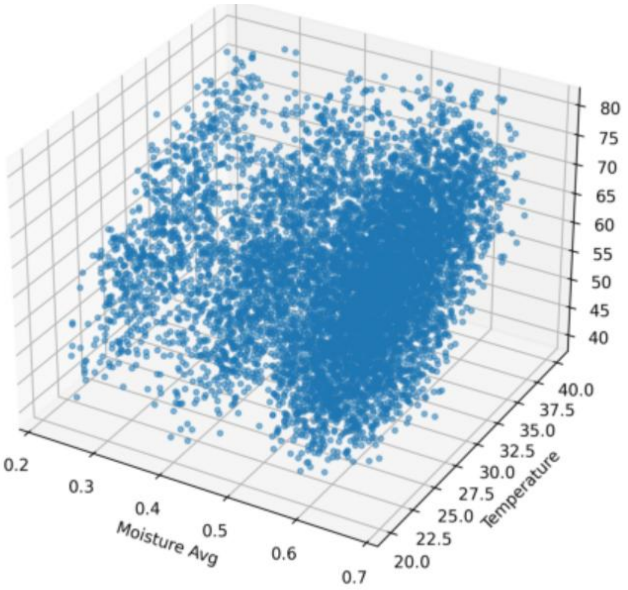
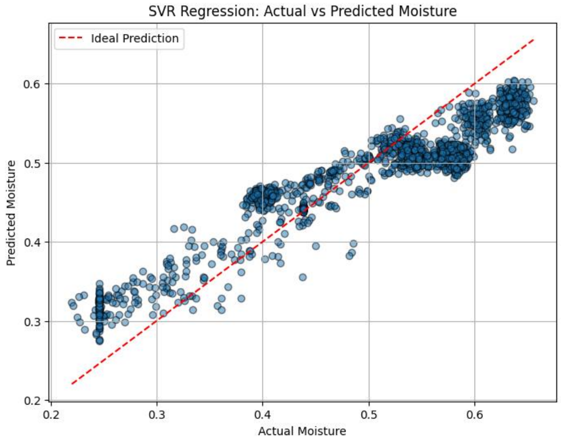


Figure 13: 3D plot of moisture, temperature, and humadity

e. SVM Classification Model

The SVR regression plots (Figure 14) highlight a powerful predictive interaction, whereby most of the moisture predictions are almost perfect matches to the ideal prediction. The predicted moisture values are for the most part between 0.30 and 0.60, whereas the actual moisture values go from around 0.25 to 0.65, thus showing a nice overlapping of the observed and the estimated values.



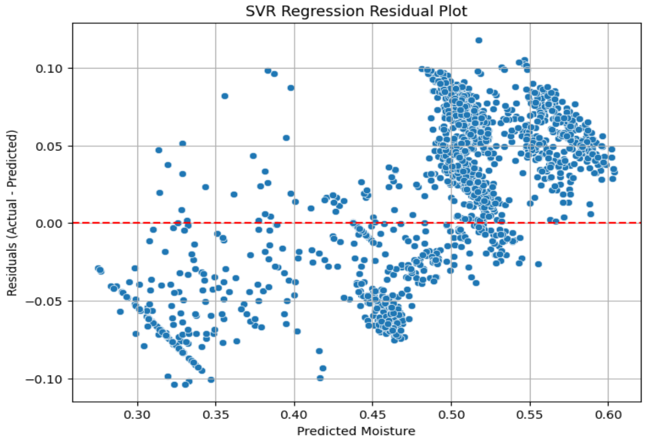
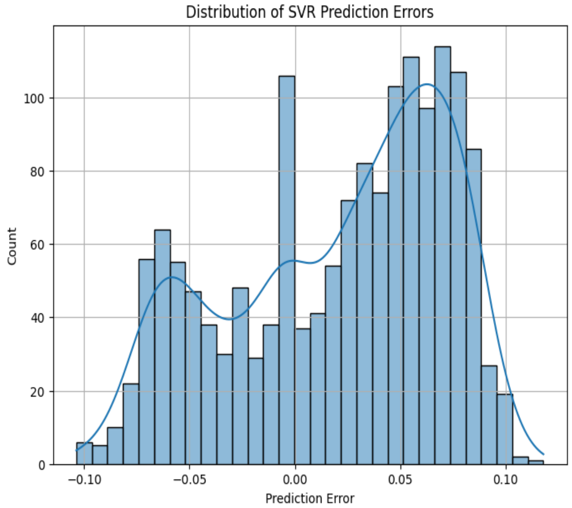


Figure 14: SVM regression: Actual vs predicted moisture

Figure 15: SVM regression residual plot.

Figure 15 shows that the model is accurate to a good extent because in it most of the residuals are between –0.05 and +0.05. There are a few points that go beyond –0.10, thus showing that the prediction is a bit lower than the actual value in the range of the medium moisture and a little bit higher at the higher values. In general, the positions and the numbers of the points and residuals show that the SVR model is working well and is able to capture the moisture trend effectively while leaving only small, localized errors.

As indicated by the left part of Figure 16, most values associated with forecast error for SVR are concentrated in a range of approximately -0.08 to +0.10; the greatest number of occurences occur at about +0.04 which suggests that there is a slight tendency to over-predict results based on this type of modelling technique. On the right side of Figure 16 there is a smaller cluster of error values at around -0.05 suggesting an occasional underestimation of results. In general, the overall distribution of forecast errors based on SVRs is well-balanced and approximately normally distributed although there is some slight right skew.



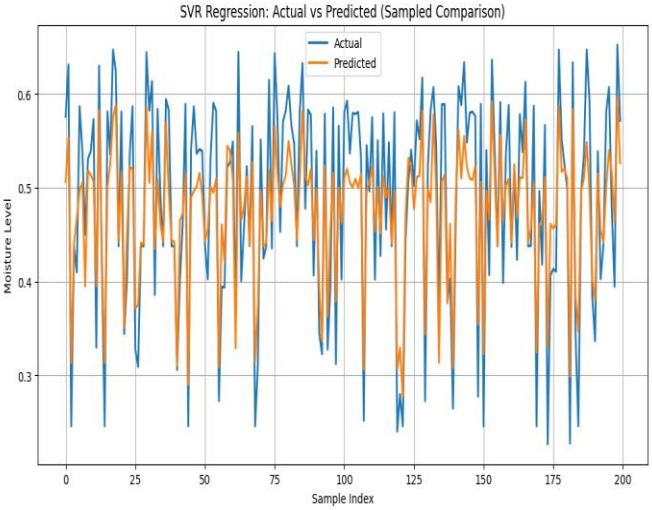
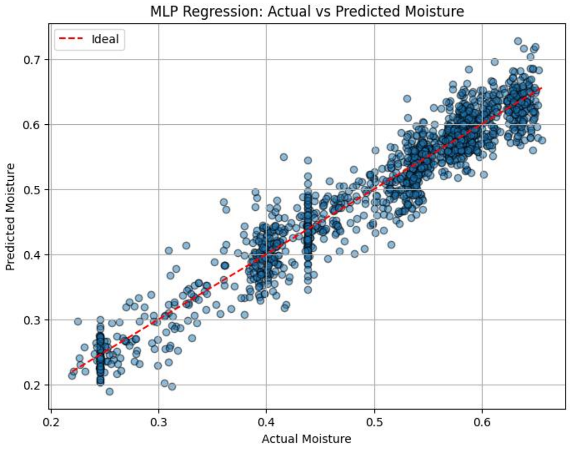


Figure 16: Distribution of SVR prediction errors **Figure 17:** SVR regression: Actual vs predicted (sample)

The sample wise comparison plot as shown in Figure 17 also indicates that the real moisture levels, in the band of roughly 0.30 to 0.65 , seem to be pretty well tracked by the forecasted values across the 200 samples. Despite slight imperfections at some peaks and dips, the fact that the two curves are aligned proves that the SVR model is always aligned to the actual moisture pattern, and predictive capabilities are always consistent even when the conditions change.

f. MLP Model Classification

The results of the MLP regression indicate that there is a high degree of correlation between actual and predicted moisture values with the majority of the points falling close to the desired reference line (as seen in Figure 18). The model effectively represents the general moisture trend especially in the intermediate area between 0.40 and 0.60, but there are minor variations at the lower values close to 0.25 and at the higher values close to 0.70.



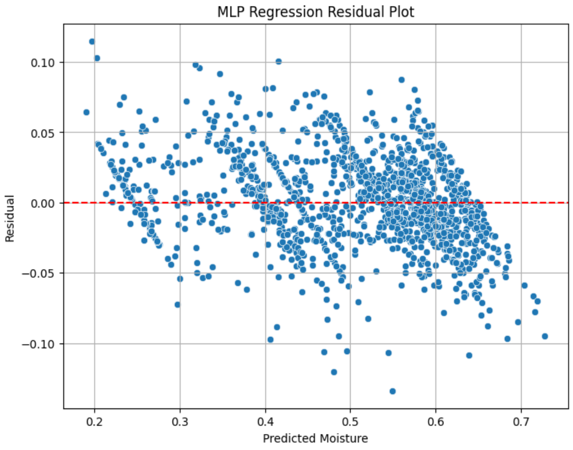


Figure 18: MLP regression: Actual vs predicted moisture **Figure 19:** MLP regression residual plot.

The residual plot (as seen in Figure 19) also shows that the majority of residuals lie between the range of -0.07 to +0.06, which implies that the error of prediction is not high in general. There is an observable trend of under-prediction at higher levels of predicted moisture, and a bit of over-prediction is observed in the lower levels. Regardless of these local variations, the residual values are fairly even about zero which proves that the MLP model provides consistent and fairly accurate moisture forecasts.

The distributions of MLP prediction error have almost normal shapes and are symmetric around zero for all the features, which suggest that the models behave in a balanced way and that there is a very small systematic bias (as seen in Figure 20). The volume of errors is between –0.04 and +0.05, while the closest to 0.00 is the maximum frequency, meaning that most of the predictions differ from the real values only to a small extent. There are a few extreme errors that reach approximately –0.11 and +0.12 but these are quite infrequent.

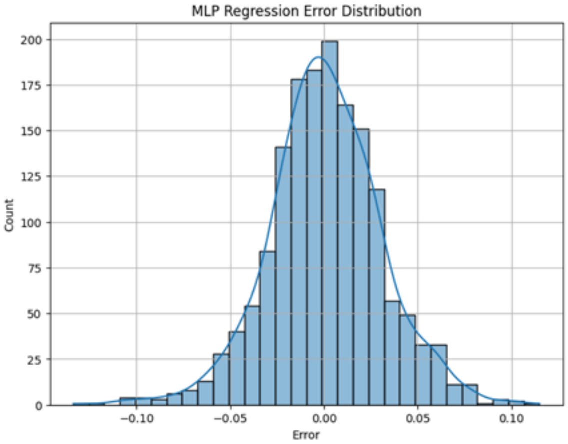


Figure 20: Distribution of MLP prediction errors

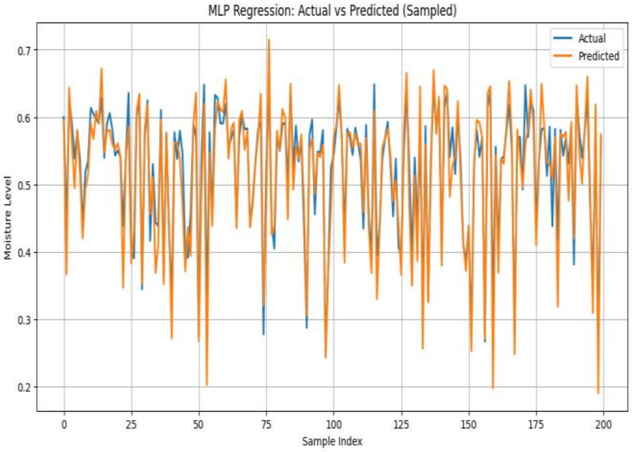


Figure 21: MLP regression: Actual vs predicted (sample)

The comparison of sampled actual vs. predicted that is shown above in figure 21, additionally, indicates that the MLP model is able to efficiently follow the change of moisture levels in 200 samples. The real values varied approximately between 0.30 and 0.65, and the predicted values also went up and down along the same line. There are some instances in which the prediction shows a sudden sharp increase, but if researchers consider the whole plot, the model is still performing steadily and accurately.

g. Hybrid Model Classification

The hybrid regression model has the highest level of precision when comparing actual vs. predicted moisture readings, which matched the expected values of all data measurements (as seen in Figure 22). The predicted moisture levels ranging between 0.25 and 0.65 were predicted with exceptional accuracy, while the aggregate distribution of the measurements showed reduced variability as compared to the separate models.

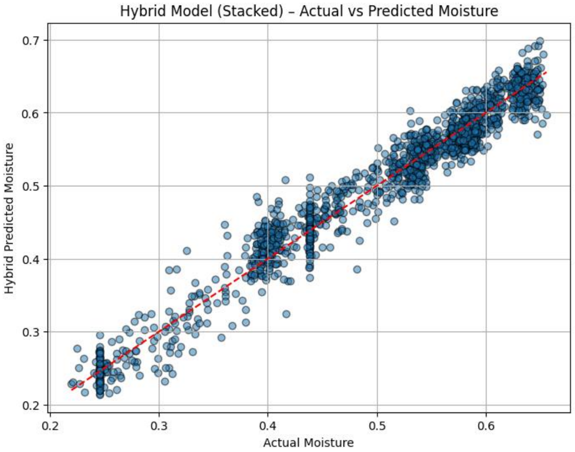


Figure 22: Hybrid regression: Actual vs predicted moisture

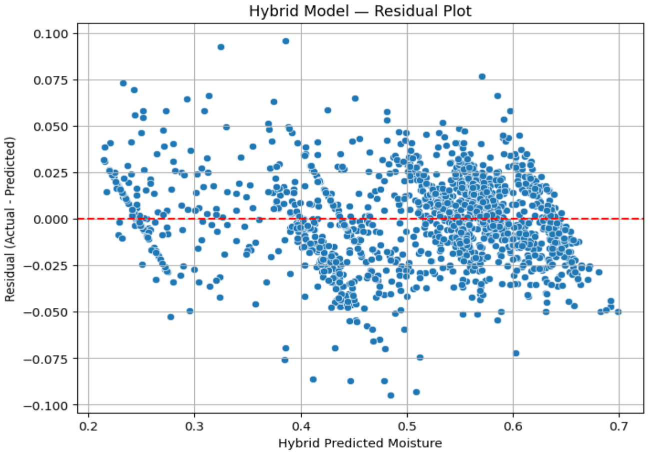


Figure 23: Hybrid regression residual plot.

Residual plot (see figure 23) also corroborates this performance, with most of the residual data concentrated around –0.04 and +0.03 which indicates a minimal error. A few outliers have values of approximately –0.10 and +0.09 but they are not common. The overall distribution does not reveal any significant systematic bias, which proves the stacked hybrid model is an effective method of improving the accuracy and stability of the prediction compared with its single models.

The error distribution of the hybrid model displays a very regular and tight distribution around zero, demonstrating good predictability stability. Most errors range between –0.03 and +0.03, with more errors in the region of very close to 0.00 which indicate very little deviation from actual moisture values. Only a few rare errors extend toward –0.09 or +0.09, demonstrating the model's strong reliability.

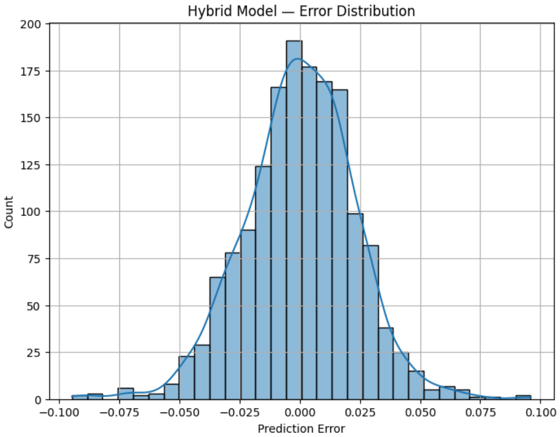


Figure 24: Distribution of Hybrid prediction errors

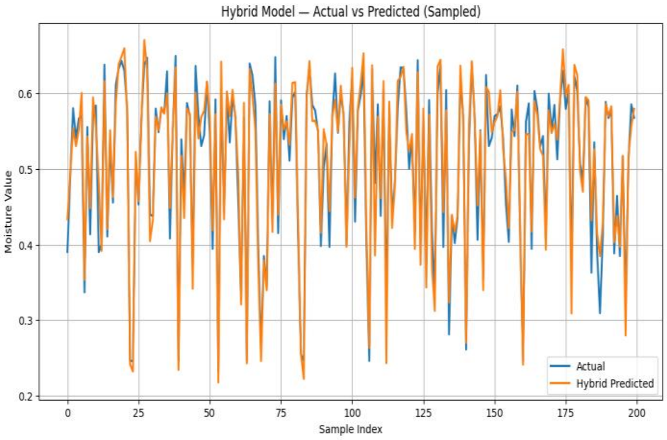


Figure 25: Hybrid regression: Actual vs predicted (sample)

The comparison between the sampled actual and predicted values as seen in Figure 25, which is less visual but more precise, is in line with this achievement of the hybrid model as it shows that the hybrid predictions are almost at the same level with the real data for most of the 200 observations, the moisture levels changing roughly from 0.35 to 0.65. While there are some sudden and significant differences that show up from time to time, the general matching is still very close, thus the hybrid model being able to demonstrate better consistency and higher accuracy than the individual SVR and MLP models.

h. Evaluation of Performance Metrics

The comparison of performance of SVM, MLP and Hybrid models as seen in Table 3 indicates an obvious increase of the prediction accuracy with the transition of classical to modern learning methods. The best error values are shown by the SVM model, its RMSE is 0.0549, and MAE is 0.0485, which means that it has a moderate predictive accuracy and the lowest R^2 of 0.7573.

Table 3: Evaluation of Performance Metrics

Model	**RMSE**	**MAE**	**MAPE (%)**	**R^2 Score**
SVM	0.05492	0.04848	10.22	0.7573
MLP	0.03027	0.02309	4.90	0.9263
Hybrid	0.02294	0.01778	3.83	0.9577

The MLP model works much better with a lower RMSE of 0.0303 and MAE of 0.0231 and a higher R2 of 0.9263. The Hybrid model has the most satisfactory performance with the lowest RMSE (0.0229), lowest MAE (0.0178), and the minimum MAPE of 3.83%. The high predictive power and high reliability are attested by its R^2 value of 0.9577. Figure 26 shows the comparison graph of model based on performance metrics.

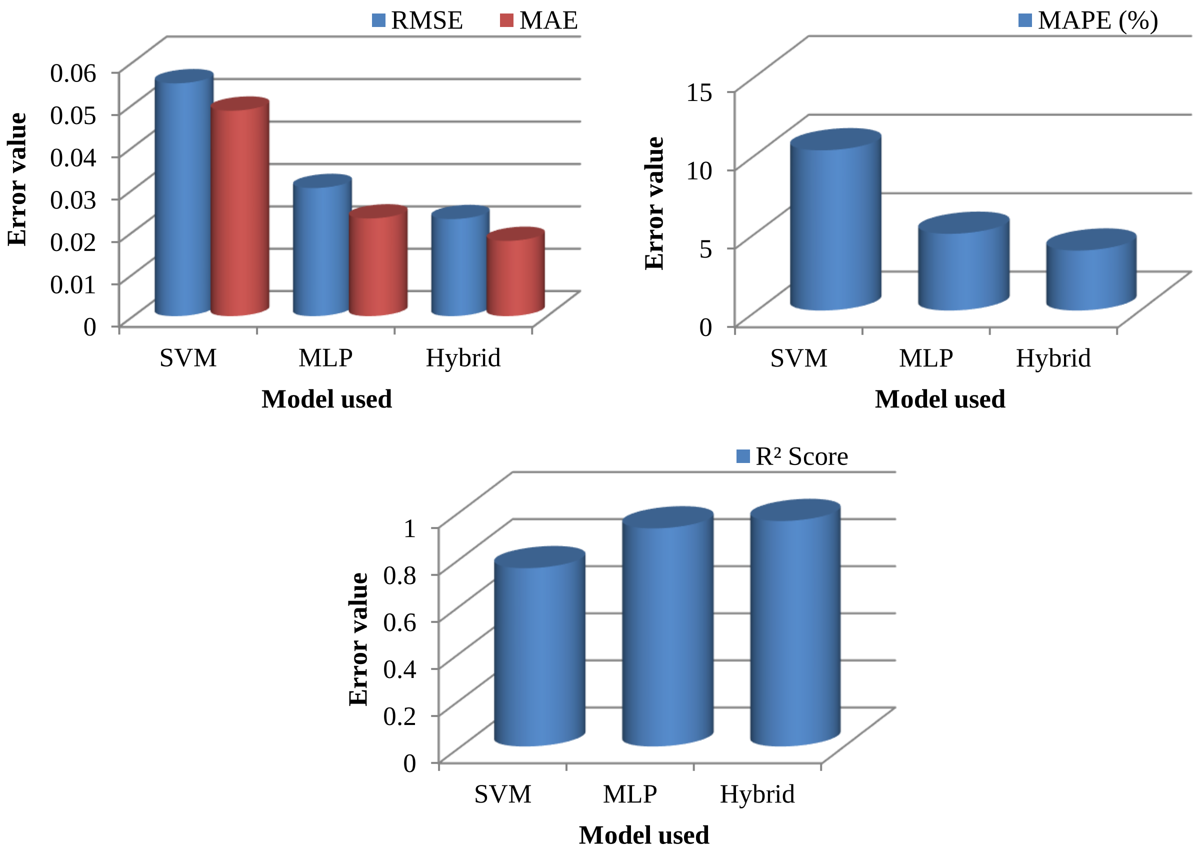


Figure 26: Comparison graph of model based on performance metrics.

The effectiveness of the comparison between the conventional and ML/IoT-optimized farming activities makes it clear that efficiency advantages are gained through the use of intelligent resource management (as seen in Figure 27). The water used under the traditional methods is reduced to approximately 1000 L whereas when using the optimized method, it consumes approximately 620 L, showing that it reduces by almost 38%. There are also less fertilizers (50 kg to approximately 35 kg) with more specific nutrient applications.

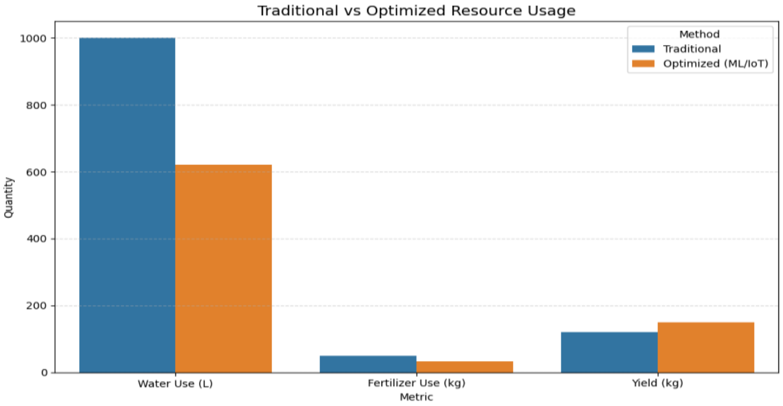


Figure 27: Traditional vs Optimized (Water, Fertilizer, Yield)

The optimized approach yields more output (150-120kg) with a reduction of approximately 25% in resources used. These findings suggest that data-driven optimization does not only save input but also improve productivity, thus making the whole system more sustainable, cost-efficient, and environmentally friendly than the traditional forms of agriculture.

The performance enhancement chart shows the big gains obtained due to the optimized ML/IoT-based solution (as seen in Figure 28). Almost 38% of water savings are made, which shows that the irrigation intensity has been reduced considerably by the accurate monitoring and forecasting of the moisture content. The reduction of fertilizer is by close to 36% that is to say that the management of nutrients has become more effective, and less wastage has been produced.

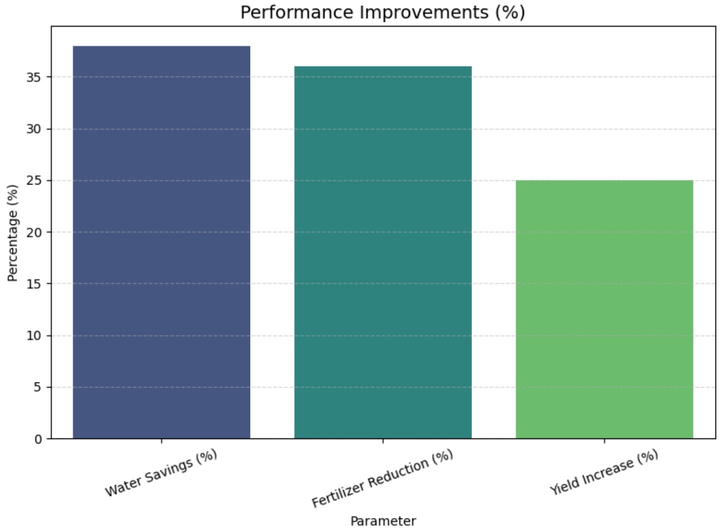


Figure 28: Bar chart for Water Savings, Fertilizer Reduction, Yield Increase

During that time, crop production rises by around 25% which shows that improved resource allocation has a direct positive impact on productivity. In total, these enhancements demonstrate that data-driven decisions can drive sustainable agriculture practices which reduce inputs and increase output, not just protecting the environment, but also turning a profit.

5. Conclusion

The study showed that the precision farming and sustainable use of resources system in the IoT research space are highly feasible. The sensors measuring NPK, moisture and temperature are real-time soil sensors, which are integrated in a networked system with the Arduino Mega microcontroller and can be analyzed via the cloud. In predictive modeling, data cleaning smoothing and outlier removal are the preprocessing steps which should be performed to get accurate and reliable datasets. Utilizing data from the Agriculture IoT 2024, machine learning models such as the multilayer perceptron (MLP), support vector machine (SVM), and hybrid of the two models were employed to predict soil moisture and nutrient requirements for various crops. The hybrid model was best among these with RMSE 0.0229 MAE 0.0178 R 0.9577 which indicates better predictive power and reliability of this system practically tested to reduce water by 38 and fertilizer by 36% while increasing crop yield dramatically by 25% indicating that IOT sensing machine learning metaverse visualization can be realized for providing actionable insights toward efficiency sustainability productivity in agriculture. In total offered framework is very strong solution that could help make decisions based on the data that facilitate precision farming more resources efficiently, less damaging to environment and able achieve maximum crop yield.

Reference

1. Mohyuddin, Ghulam, Muhammad Adnan Khan, Abdul Haseeb, Shahzadi Mahpara, Muhammad Waseem, and Ahmed Mohammed Saleh. "Evaluation of machine learning approaches for precision farming in smart agriculture system: a comprehensive review." *IEEE access* 12 (2024): 60155-60184.
2. Shaikh, Tawseef Ayoub, Waseem Ahmad Mir, Tabasum Rasool, and Shabir Sofi. "Machine learning for smart agriculture and precision farming: towards making the fields talk." *Archives of Computational Methods in Engineering* 29, no. 7 (2022): 4557-4597.
3. Karunathilake, E. M. B. M., Anh Tuan Le, Seong Heo, Yong Suk Chung, and Sheikh Mansoor. "The path to smart farming: Innovations and opportunities in precision agriculture." *Agriculture* 13, no. 8 (2023): 1593.

4. Sharma, Abhinav, Arpit Jain, Prateek Gupta, and Vinay Chowdary. "Machine learning applications for precision agriculture: A comprehensive review." *IEEe Access* 9 (2020): 4843-4873.
5. Mekonnen, Yemeserach, Srikanth Namuduri, Lamar Burton, Arif Sarwat, and Shekhar Bhansali. "Machine learning techniques in wireless sensor network based precision agriculture." *Journal of the Electrochemical Society* 167, no. 3 (2019): 037522.
6. Dhanaraju, Muthumanickam, Poongodi Chenniappan, Kumaraperumal Ramalingam, Sellaperumal Pazhanivelan, and Ragunath Kaliaperumal. "Smart farming: Internet of Things (IoT)-based sustainable agriculture." *Agriculture* 12, no. 10 (2022): 1745.
7. Singh, Ritesh Kumar, Rafael Berkvens, and Maarten Weyn. "AgriFusion: An architecture for IoT and emerging technologies based on a precision agriculture survey." *IEEE Access* 9 (2021): 136253-136283.
8. Abioye, Emmanuel Abiodun, Oliver Hensel, Travis J. Esau, Olakunle Elijah, Mohamad Shukri Zainal Abidin, Ajibade Sylvester Ayobami, Omosun Yerima, and Abozar Nasirahmadi. "Precision irrigation management using machine learning and digital farming solutions." *AgriEngineering* 4, no. 1 (2022): 70-103.
9. Woźniak, Marcin, and Muhammad Fazal Ijaz. "Recent advances in big data, machine, and deep learning for precision agriculture." *Frontiers in Plant Science* 15 (2024): 1367538.
10. Haval, Abhijeet Madhukar, and F. Rahman. "Application of machine learning techniques and the Internet of Things for smart, sustainable agriculture." In *BIO Web of Conferences*, vol. 82, p. 05021. EDP Sciences, 2024.
11. Awais, Muhammad, Xiuquan Wang, Sajjad Hussain, Farhan Aziz, and Muhammad Qasim Mahmood. "Advancing precision agriculture through digital twins and smart farming technologies: a review." *AgriEngineering* 7, no. 5 (2025): 137.
12. Nasirahmadi, Abozar, and Oliver Hensel. "Toward the next generation of digitalization in agriculture based on digital twin paradigm." *Sensors* 22, no. 2 (2022): 498.
13. Dawn, Nabarun, Souptik Ghosh, Aloke Saha, Soumya Chatterjee, Tania Ghosh, Sagnik Guha, Subhajit Sarkar, Pronoy Mukherjee, and Tanmay Sanyal. "A review on digital twins technology: A new frontier in agriculture." In *Artificial Intelligence and Applications*, vol. 2, no. 4, pp. 250-262. 2024.
14. Quy, Vu Khanh, Nguyen Van Hau, Dang Van Anh, Nguyen Minh Quy, Nguyen Tien Ban, Stefania Lanza, Giovanni Randazzo, and Anselme Muzirafuti. "IoT-enabled smart agriculture: architecture, applications, and challenges." *Applied Sciences* 12, no. 7 (2022): 3396.
15. Gund, Rajesh, Chetan M. Badgujar, Sathishkumar Samiappan, and Sindhu Jagadamma. "Application of Digital Twin Technology in Smart Agriculture: A Bibliometric Review." *Agriculture* 15, no. 17 (2025): 1799.
16. Ali, Zain Anwar, Mahreen Zain, Raza Hasan, Hussain Al Salman, Bader Fahad Alkhamees, and Faisal Abdulaziz Almisned. "Circular economy advances with artificial intelligence and digital twin: Multiple-case study of Chinese industries in agriculture." *Journal of the Knowledge Economy* 16, no. 1 (2025): 2192-2228.
17. Ametefe, Divine Senanu, Norhayati Hussin, Dah John, George Dzorgbenya Ametefe, Abdulmalik Adozuka Aliu, and Zainab Abdi Ali. "Revolutionising agriculture for food security and environmental sustainability: A perspective on the role of digital twin technology." *CABI Reviews* 19, no. 1 (2024).
18. Sharma, Amit, Ashutosh Sharma, Alexey Tselykh, Alexander Bozhenyuk, Tanupriya Choudhury, Madani Abdu Alomar, and Manuel Sánchez-Chero. "Artificial intelligence and internet of things oriented sustainable precision farming: Towards modern agriculture." *Open Life Sciences* 18, no. 1 (2023): 20220713
19. Vatin, Nikolai Ivanovich, Sanjeev Kumar Joshi, Puja Acharya, Rajat Sharma, and N. Rajasekhar. "Precision agriculture and sustainable yields: insights from IoT-driven farming and the precision agriculture test." In *BIO conferences*, vol. 86, p. 01091. EDP Sciences, 2024.
20. Jin, Xue-Bo, Xing-Hong Yu, Xiao-Yi Wang, Yu-Ting Bai, Ting-Li Su, and Jian-Lei Kong. "Deep learning predictor for sustainable precision agriculture based on internet of things system." *Sustainability* 12, no. 4 (2020): 1433.
21. Panduranga, Kiran Muniswamy, and Roopashree Hejjaji Ranganathasharma. "Sustainability insights on learning-based approaches in precision agriculture in internet-of-things." *International Journal of Electrical and Computer Engineering (IJECE)* 14, no. 3 (2024): 3495-3511.
22. Ferrández-Pastor, Francisco Javier, Juan Manuel García-Chamizo, Mario Nieto-Hidalgo, and José Mora-Martínez. "Precision agriculture design method using a distributed computing architecture on internet of things context." *Sensors* 18, no. 6 (2018): 1731.
23. Bakthavatchalam, Kalaiselvi, Balaguru Karthik, Vijayan Thiruvengadam, Sriram Muthal, Deepa Jose, Ketan Kotecha, and Vijayakumar Varadarajan. "IoT framework for measurement and precision agriculture: predicting the crop using machine learning algorithms." *Technologies* 10, no. 1 (2022): 13.
24. Hemal, Md Moniruzzaman, Suman Saha, and Kamruddin Nur. "An IoT and Machine Learning-driven Advanced Greenhouse Farming System for Precision Agriculture." *International Journal of Computing* 18, no. 1 (2025): 1-17.

25. Omer, Batool Anwar, Mohamed Mabrouk Morsey, Islam Hegazy, Zaki Taha Fayed, and Taha El-Arif. "Toward Precision Agriculture: Integrating Machine Learning Techniques for Smart Farming Systems." *IEEE Access* 12 (2024): 172910-172922.

26. Thilakarathne, Navod Neranjan, Muhammad Saifullah Abu Bakar, Pg Emerolylariffion Abas, and Hayati Yassin. "A cloud enabled crop recommendation platform for machine learning-driven precision farming." *Sensors* 22, no. 16 (2022): 6299.

27. Wongchai, Anupong, Surendra Kumar Shukla, Mohammed Altaf Ahmed, Ulaganathan Sakthi, and Mukta Jagdish. "Artificial intelligence-enabled soft sensor and internet of things for sustainable agriculture using ensemble deep learning architecture." *Computers and Electrical Engineering* 102 (2022): 108128.

28. https://www.kaggle.com/datasets/wisam1985/iot-agriculture-2024

29. Kamel, Islam Mohamed, and Hamid El Bilali. "Sustainability transition to organic agriculture through the lens of the multi-level perspective: case of Egypt." *Organic Agriculture* 12, no. 2 (2022): 191-212.

30. Boniecki, Piotr, Maciej Zaborowicz, Agnieszka Pilarska, and Hanna Piekarska-Boniecka. "Identification process of selected graphic features apple tree pests by neural models type MLP, RBF and DNN." *Agriculture* 10, no. 6 (2020): 218.

31. Olier, Ivan, Annabel Sansom, Paulo Lisboa, and Sandra Ortega-Martorell. "Using MLP partial responses to explain in-hospital mortality in ICU." In *2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, pp. 1-5. IEEE, 2020.

32. Taghizadeh, Sepideh, Hossin Navid, Reza Adiban, and Yasser Maghsodi. "Harvest chronological planning using a method based on satellite-derived vegetation indices and artificial neural networks." *Spanish Journal of Agricultural Research* 17, no. 3 (2019): e0206-e0206.

33. Nosratabadi, Saeed, Sina Ardabili, Zoltan Lakner, Csaba Mako, and Amir Mosavi. "Prediction of food production using machine learning algorithms of multilayer perceptron and ANFIS." *Agriculture* 11, no. 5 (2021): 408.

34. Aparecido, Lucas Eduardo de Oliveira, José Reinaldo da Silva Cabral de Moraes, Glauco de Souza Rolim, Lucieta Guerreiro Martorano, Sabrina dos Santos Soares, Kamila Cunha de Meneses, Cicero Teixeira Silva Costa et al. "Neural networks in spatialization of meteorological elements and their application in the climatic agricultural zoning of bamboo." *International Journal of Biometeorology* 62, no. 11 (2018): 1955-1962.

35. Islam, Sumaiya, Samsuzzaman, Md Nasim Reza, Kyu-Ho Lee, Shahriar Ahmed, Yeon Jin Cho, Dong Hee Noh, and Sun-Ok Chung. "Image processing and support vector machine (SVM) for classifying environmental stress symptoms of pepper seedlings grown in a plant factory." *Agronomy* 14, no. 9 (2024): 2043.

36. Ayalew, Abebe Temesgen, and Tarun Kumar Lohani. "Prediction of crop yield by support vector machine coupled with deep learning algorithm procedures in Lower Kulfo watershed of Ethiopia." *Journal of Engineering* 2023, no. 1 (2023): 6675523.

37. Sharma, Shivi, Geetanjali Rathee, and Hemraj Saini. "Big data analytics for crop prediction mode using optimization technique." In *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, pp. 760-764. IEEE, 2018.

38. Kumar, G. Kranthi, Manoj L. Bangare, Pushpa M. Bangare, Chanda Raj Kumar, Roop Raj, José Luis Arias-Gonzáles, Batyrkhan Omarov, and Md Solaiman Mia. "Internet of things sensors and support vector machine integrated intelligent irrigation system for agriculture industry." *Discover Sustainability* 5, no. 1 (2024): 6.

39. Bhattacharyya, Debnath, Eali Stephen Neal Joshua, N. Thirupathi Rao, and Tai-hoon Kim. "Hybrid CNN-SVM classifier approaches to process semi-structured data in sugarcane yield forecasting production." *Agronomy* 13, no. 4 (2023): 1169.