VER
Vascular & Endovascular Review

Real-Time Prediction of Crop Diseases Using IoT-Enabled Data Acquisition and Machine Learning

Ms. Susmita Arun Meshram[1], Dr. Nitin K. Choudhari[2]

[1]Research Scholar at Post Graduate Teaching Department of PGTD
Computer Science & Electronics, Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur.
susmitameshram0505@gmail.com
[2]Guide & Principal Priyadarshini Bhagwati College of Engineering Nagpur (MS) Rashtrasant Tukadoji Maharaj Nagpur University
drnitinchoudhari@gmail.com

ABSTRACT

Crop diseases are a major threat to global agriculture by affecting yield quality and quantity while leading to economic impacts and food insecurity. Risk management measures must be initiated with few steps for early detection and immediate treatment or control in order to mitigate the potential consequences of crop disease. This paper presents a unified proposal for the first real-time prediction of crop diseases using Internet of Things (IoT)-enabled data acquisition fused with machine learning (ML). The proposal is based on the system and includes a collection of IoT devices including environmental sensors, smart-cameras, drones, low-power wide-area networks (LPWANs) for transmitting the data that collect and relay environmental information (e.g., temperature, humidity, soil moisture, plant wellbeing, and other information). Furthermore, all field-oriented data use, relatively speaking, better, centralized cloud storage (and processing) which ensures the data collected from IoT data transmissions is accessible. Preprocessing techniques that cleaned and normalized the data were utilized to enhance the value of the inputs before uploading a data stream. Data would then be fed into a Convolutional Neural Network (CNN) to capture high-resolution images of the leaves in order to assess disease symptoms, compared to other environmental and agronomic data, respectively. Environmental/ agronomic data would provide access to supervised learning models (e.g., Random Forest and Gradient Boosting) for assessing or detecting possible early-stage disease patterns.

KEYWORDS: Crop Disease Prediction; Internet of Things (IoT); Machine Learning (ML); Convolutional Neural Networks (CNN); Random Forest; Gradient Boosting; Smart Agriculture; Precision Farming; Environmental Sensors; Real-Time Monitoring; Disease Detection; Sustainable Agriculture.

How to Cite: Susmita Arun Meshram, Nitin K. Choudhari, (2025) Real-Time Prediction of Crop Diseases Using IoT-Enabled Data Acquisition and Machine Learning, Vascular and Endovascular Review, Vol.8, No.12s, 334-346.

INTRODUCTION

The current approaches for identifying crop diseases rely heavily on the visual observations and human monitoring of agricultural experts. Delays in diagnosis and errors due to human error characterize these procedures, which are both labor-intensive and time-consuming. Crops in less developed or far-flung regions are more likely to be susceptible to illnesses that go undetected and spread rapidly because farmers in these regions lack access to agricultural experts and timely information sources[1][2]. The need for a smart, real-time crop health surveillance system that can accurately and efficiently forecast disease outbreaks and detect them in the field is growing. Possible answers include the expanding realm of digital agriculture, new possibilities in the Internet of Things (IoT) sector, and the constantly evolving discipline of machine learning (ML). Temperature, humidity, soil moisture, leaf color, and air conditions are just a few of the environmental and plant characteristics that are currently being measured by sensors that are internet of things (IoT) enabled[3]. Combining the data collected by IoT sensors with top-notch ML algorithms allows for real-time analysis, which may spot irregularities and foretell possible disease outbreaks before they spiral out of control. Approximately 58% of India's population is directly or indirectly involved in agriculture. Unfortunately, conventional agricultural practices cannot produce enough food to meet the needs of the Indian population, who need a varied and nutrient-dense diet. As a means of meeting the increasing need[4].

The fertility of the soil, crop quality, and yield all significantly decline as a result of ignorance, overdose of fertilization, misinformation, and improper farming techniques. IoT framework with machine learning (ML) for soil and crop analysis is essential in contemporary agriculture as it facilitates accurate, efficient, and sustainable agricultural methods. Recently, ML has made its way into the realm of trans disciplinary agri-technologies, creating new possibilities for data-intensive research through the integration of big data analytics with high-performance computers[5][6]. It involves the use of revolutionary technologies, such as IoT and Artificial Intelligence (AI) with deep learning. In addition to pertinent sensors and camera modules, to gather data on major soil nutrients and images of crop plants for additional analysis to predict diseases. This made it possible to understand the condition of the soil and crop disease. It provides farmers with up-to-date information on how to apply fertilizer and pesticides to their current and future crops in a way that protects consumer health while increasing crop quality and yield[7][8].

The initial and important stage in agricultural development is the selection of the appropriate crop to optimize productivity. An essential factor in increasing production is the careful selection of crops depending on the unique properties of the soil[9][10][11]. To protect the agricultural region, it is important to conduct tests and determine the required fertilizer. In India, soil testing was first conducted in 1955–1956, with the Indian Agricultural Research Institute (IARI) serving as the central office for liaising with all other soil testing facilities around the nation. The laboratory provides farmers and other clients with soil, plant, manure, and irrigation water analysis services. Before choosing any crop, the quality of the soil must also be assessed for minerals, pH, and other factors. Soybeans are an exceptionally adaptable and lucrative crop that has extensive effects on nutrition, agriculture, industry, and the economy.

Their capacity to supply vital nutrients, promote sustainable agricultural methods, and contribute to a wide range of industrial uses highlights their significance in present and future food and industrial systems. Soybean diseases have a substantial impact on the productivity and quality of crops, resulting in financial losses for farmers and influencing the global supply of food and animal feed. Bureau of Indian Standards [6] states that the United States, Brazil, Argentina, China, and India are the five main producers of soy in the world, making up 90% of global soy production. Brazil, the United States, Argentina, and Paraguay are the main exporters of oil, soybean meal, and soybeans. In 2017, the combined share of these four nations in the world's soy product commerce was over 91%. In recent years, the soybean harvest in India has consistently maintained its position as the leading crop among the nine field oilseeds cultivated in the country. The geographical regions of Madhya Pradesh, Maharashtra, and Rajasthan collectively constitute the predominant soybean cultivation areas in Central India. The other periphery states of Chhattisgarh, Karnataka, and Telangana do the same[2]

The IoT framework developed in this research comprises several components, including soil sensors, IoT module for data processing, cloud-based machine learning analysis, and a mobile interface for farmers. Utilizing the central hub of an Arduino Uno and ESP 8266 to collect and transmit information such as soil moisture, pH, and nitrogen levels to the cloud and then machine learning models analyze the data to produce recommendations; the recommendations delivered to farmers' mobile devices will enable them to make decisions about fertilizer, disease and crop health[4][7]. The framework provides real time soil health monitoring and early diagnosis of crop diseases for enhanced crop management and productivity. The systems adaptability should enable scalability for other crops and farming situations, where it could continue to benefit farmers beyond just soybean production. Furthermore, this research has a positive environmental impact, as it will reduce unneeded fertilizer and pesticide applications contributing to cost savings for farmers while also preserving soil health. More broadly, this research strengthens the livelihood of farmers through improved yield quality and quantity. The research offers a scalable and flexible framework that encourages future precision agriculture developments by leveraging IoT, AI, and machine learning that can innovate in agriculture. Our results indicate the potential for more precision farming applications using artificial intelligence and to investigate the framework applications for other crops and potentially more efficiency and sustainability in farming practices around the world[8][9][10][11].

In addition, the study offers a CNN model that processes soybean leaf images to detect disease earlier. This automated disease detection system, along with real-time soil monitoring, enables farmers to take preventative measures before a disease spreads, saving the crop and limiting crop-loss. The agricultural industry currently lacks a viable, scalable automated system that can predict crop disease in real-time from the field using field data[7]. Existing systems have use limitations (e.g. static forecasting systems that use pre-recorded datasets) or cannot be considered real-time due to data processing capabilities (network or computing). In general, the initial cost of implementing a predictive system or solution can be seen as too large of a venture for the agricultural industry, in being able to pay for loss insurance or providing scaling. Additionally, poor network bandwidth in rural areas does not offer an easy solution to if connectivity evaporates or is interrupted in areas where smart-connected devices may not have other capabilities to "acquire" data[4]. This paper addresses this gap of needing a viable and scalable prediction system that uses IoT enabled data acquisition capabilities and various machine learning algorithms.

LITERATURE REVIEW

Reference	Year	Focus Area	Technologies/Models Used	Key Findings/Contributions	Performance Metrics
Md Abrar Jahin et al.	2025	Soybean Leaf Disease Detection	Hybrid Sequential CNN-Graph Neural Network (GNN) framework (MobileNetV2, GraphSAGE), Grad-CAM, Eigen-CAM	Proposed an interpretable hybrid model addressing visually similar symptoms and inter-image relational dependencies. Achieved cross-modal interpretability and computational efficiency.	Accuracy: 97.16% (surpassing standalone CNNs $\le$95.04% and traditional ML $\le$77.05%).
Nilesh B. Korade et al.	2025	Soybean Disease Detection	Deep Learning (AlexNet, VGG-16, Inception-v3, EfficientNetV2B0, ResNet50), YOLO for leaf extraction	ResNet-50 outperformed other models in predicting soybean conditions from captured images, showing robustness across all classification metrics.	ResNet-50 most accurate (specific metrics not provided but stated as highest across all).

Zainab A. Abdulazeez et al.	2025	Soybean Disease Prediction (Healthy vs. Infected)	Machine Learning (Support Vector Machine, Random Forest), Feature Selection (Gain Ratio, Correlation), Cross-validation	Developed a smart forecasting model to differentiate between healthy and infected soybean plants. Performance metrics are impacted by reducing soybean traits.	Accuracy, F-measure, Specificity, Executing Time, Sensitivity (
Ashish Saini et al.	2025	Plant Disease Categorization in IoT Network	IoT, Henry Gas Chicken Swarm Optimization (HGCSO), Deep Residual Network (DRN), Caviar Henry Gas Chicken Swarm Optimization (CHGCSO), Median Filtering, HoG, Statistical Features, SLIF, LTP	Developed a new model for plant disease categorization within an IoT network. CHGCSO-based DRN outperformed existing methods.	Accuracy: 94.3%, Sensitivity: 93.3%, Specificity: 92%, F1-score: 93%.
Ritesh Maurya et al.	2025	Tomato Plant Leaf Disease Detection	Deep Learning (RAI-Net: Residual-Attention-Inception-Net), Fine-tuned ResNet18, Inception module, Channel Attention, Grad-CAM	Proposed RAI-Net model for effective feature extraction and multi-scale feature analysis. Improved interpretability with Grad-CAM.	Accuracy: 97.88% (on test set of 4595 images across 10 classes).
Muhammad Shoaib et al.	2025	Review: Deep Learning for Plant Disease and Pest Detection	Deep Learning (Classification, Detection, Segmentation Networks)	Detailed analysis showing how deep learning models outperform more conventional approaches. Stresses how AI might revolutionize agricultural diagnosis.	Classification: >95% accuracy; Detection/Segmentation: >90% precision.
Manju G et al.	2024	Crop Prediction based on Soil Fertility	Machine Learning (k-nearest neighbors (KNN)), Real-time Soil Fertility Analyzer	Developed a real-time soil fertility analyzer for crop prediction. KNN showed the highest performance for predicting suitable crops (coconut, ginger, plantain, tapioca).	Accuracy: 84%, Precision: 85%, Recall: 88.8%, Specificity: 92.4%.
Bong-Hyun Kim et al.	2024	Soybean Leaf Disease Detection	Machine Learning-Convolutional Neural Network (CNN)	Created a Smart Detection System that can accurately identify illnesses affecting soybean leaves. Convolutional neural network (CNN) model based on data from Diabrotica Speciosa, Caterpillar, and Healthy soybean leaves.	Accuracy: 95%.
Letícia Bernabé Santos et al.	2024	Soybean Yield Prediction	Random Forests Algorithm, Linear Regression	Investigated the potential of random forests as a forecasting tool using information from crop management systems. When compared to more conventional statistical methods, random forest.	Identified optimal hyperparameters for random forests and a data threshold for optimal results (specific metrics not provided).

Gunaganti Sravanthi et al.	2024	Crop Recommendation & Disease Prediction (IoT-based)	IoT, Multi-level Kronecker Guided Pelican Convolutional Neural Network (MKGPCNN), Combined Graph Sample and Aggregate Attention Network (CGSAAN)	An Internet of Things (IoT) framework was introduced with the objective of improving precision farming via crop suggestion and disease diagnosis..	Crop Recommendation: Accuracy: 99%, Precision: 99.5%, Recall: 99.6%. Disease Identification: Accuracy: 98%, Precision: 99%, Recall: 98.9%.
Payam Delfani et al.	2024	Review: Plant Disease Forecasting Models	IoT, Machine Learning, Artificial Intelligence	Describes the function of illness forecasting models, with an emphasis on recent developments and AI/ML applications. Explores current state of open-source AI models and potential future developments towards more transparent and interpretable models.	N/A (Review Paper)
Martin Kuraduseng e et al.	2024	Crop Yield Prediction (Irish Potato & Maize)	IoT, Machine Learning	Using both past and present meteorological conditions as well as data on past and present crop yields, we created a system that uses IoT and ML to forecast future harvests.	Mean Absolute Percentage Error (MAPE): 0.339, 0.309 (Irish Potato), 0.177 (Maize).
Zaiba Khan et al.	2024	Review: IoT in Precision Agriculture (PA)	Internet of Things (IoT), Big Data	Comprehensive review of current innovative means of utilizing IoT in PA practices, discussing design architecture of communicating systems and technologies.	N/A (Review Paper)
Abbas Jafar et al.	2024	Review: Automated Leaf Disease Diagnosis (Tomato, Chilli, Potato, Cucumber)	AI, IoT Sensors, Machine Learning, Deep Learning	Identifies four agricultural diseases, describes their symptoms, and outlines the techniques to anticipate these illnesses using AI. Explores model and dataset topics related to ML/DL detection.	N/A (Review Paper)
Sangyeon Lee et al.	2023	Crop Disease Risk Prediction (Strawberry, Pepper, Grape, Tomato, Paprika)	Deep Learning, Growth Environment Data (Air Temperature, Relative Humidity, Dew Point, CO2 Concentration)	Proposed a model predicting disease risk scores based on previous growth environment information.	Average AUROC: 0.917.

METHODOLOGY

Improving crop health monitoring and management via real-time disease prediction utilizing IoT-based data capture and machine learning algorithms is a groundbreaking approach in contemporary agriculture. With the help of Internet of Things (IoT) sensors spread out across fields, this system can gather data in real time and increase field-level observations without emissions while also measuring environmental and soil parameters like temperature, humidity, soil moisture, pH, and nutrient levels[12][13]. In order to conduct real-time monitoring, all of the data collected by the distributed sensors is sent to the cloud for processing. Machine learning algorithms carry out continuous processing in these cloud databases, finding and evaluating data for patterns linked to disease outbreaks. In order to forecast the occurrence of various agricultural diseases, machine learning models like decision trees, support vector machines, and deep neural networks would be trained using data that is both historical and close to real-time. Machine learning algorithms can detect subtle changes in plant physiology and environmental factors long before disease symptoms appear, allowing for the prediction of when diseases will strike without affecting harvest production[14]. Not only may machine learning and the internet of things help us spot disease outbreaks sooner, but they could also lend credence to food-specific agricultural practices that use just the amount of fertilizers and pesticides really required, leading to more efficient farming overall. Insight and actionable notifications enable farmers to react quickly using mobile applications and user-friendly dashboards.Sensors, algorithms, and, most significantly, the scalability of these technologies to handle a wide variety of crops and agricultural settings are major areas of study and development. To better comprehend plant-pathogen interactions and, more crucially, to develop solutions tailored to the unique characteristics of each crop and farming setting, agronomists, data scientists, and tech developers must collaborate[15]. For smallholder and resource-poor farmers to embrace innovations, it is crucial that IoT devices be both affordable and easily accessible. To assist desensitize these innovative areas and the traditional farming community, support educational programs in addition to conventional training programs may aid in the acceptance of these diverse new technologies in traditional farming. To conclude, the use of machine learning and data collected via the Internet of Things may improve the capacity to forecast and control agricultural diseases, making agriculture throughout the world more resilient and long-lasting.[16]

ML based smart irrigation techniques using IoT

The modernization of irrigation methods is brought along with the rise and growth of IoT and ML techniques. In the ML based irrigation models, the moisture contained in the soil will be monitored with the help of soil moisture sensors for measuring the amount of water in the soil. Upon receiving data from the soil moisture sensors, the ML algorithm will be able to decide whether irrigation is needed or not, thereby reducing the wastage of water . This decision is usually accomplished by the ML algorithmic scheme because of its learning on the data on which it is trained[17]. The process of learning is basically categorized into two groups, like unsupervised learning and supervised learning . The IoT-based irrigation module employs both the supervised as well as unsupervised learning processes . But, supervised learning was the most adopted technique and it utilizes the data which is labeled for training. Here, labeled data means that some of the inputs are already mapped to the output. Thus, supervised ML algorithmic model executes on the basis of supervision[18][19].

DL based smart irrigation techniques using IoT

A subset of ML algorithm is DL, which has the capability to learn by itself without the need of human intervention. DL algorithm is motivated by the composition and functioning of the human brain called ANN. DL is a neural network comprising of input, output, and hidden layers. The input layer is where the data is being given to the network, and the input data will be processed at the hidden layer to make predictions, detections or classifications at the output layer. The neural network can be built with one or more number of hidden layers. Even though a neural network with only one layer can produce fairly accurate prediction, addition of hidden layers can make accurate predictions[20][21].Deep Neural Networks (DNNs) contains many layers of interconnected nodes. All nodes act upon the prior layers for optimizing predictions. This process is called as forward propagation. One more process known as back-propagation utilizes algorithms, such as gradient descent, for calculating the errors occurring in predictions and later adjusts the weights and biases of the function by updating the parameters to train the neural network Forward propagation and back-propagation together make predictions and correct the errors respectively. When both these processes are performed for a period, the neural network becomes more accurate in predictions. DL is used in a variety of applications and their application in the domain of agriculture is no wonder. Smart irrigation is one such domain in which DL plays a vital role by determining whether water is needed for crop or not. Various types of DL models utilized for smart irrigation purpose include Gated Recurrent Unit (GRU), Convolutional Neural Network (CNN), LSTM, etc[22].

Efficient resource allocation in IoT environment is a challenging issue for using IoT systems in fog computing. Nowadays, node deployment in fog computing is a growing technology for upcoming IoT networks. The requirements are not satisfied at global level in existing methods. For delay-controlled communication in agricultural field, various optimal node deployment strategies are used in this study. Low mass density fog nodes are appropriate for improving communication seamlessness. So, Variable Tangent Search (VTS) algorithm is established to change the location throughput of nodes in every iteration. This method enables the node deployment at any place in the search space within the boundary. This presented algorithm is implemented and tested in MATLAB working environment[2]. By comparing to other deployment approaches, the presented approach has improved latency, better outperforms and lowering cost. In this model, established algorithms convergence speed is increased. Outcomes of this model is better than various other optimization algorithms.Internet of Things (IoT) that links physical devices to the large exchange of data in the network. Deployment of nodes are used for communication and to gather information of nodes in sensor network. The performance of multi-tier nodes in sensors are improved using node deployment in IoT-based fog environment. Various technologies of IoT has been progressively combined with other features such as cloud computing, Wireless Sensor Networks (WSN) and embedded systems in current years for the improvement of its efficiency, network throughput, load management, security and better utilization of resources [8]. Selecting an appropriate node deployment procedure will decrease the over-all cost of configuration in network. Minimum no. of sensor nodes is required for deployment and to confirm desired connectivity and coverage. Issues in node deployment can be solved effectively using various meta-heuristic methods[9][10]. Significantly, node deployment impacts connectivity, lifetime, deployment cost, data latency, connectivity and complete coverage of sensor networks [1]. Complex computing issues will occur due to the increased no. of devices and various data generation. Calculation is given to the end client and this approach in cloud environment is not adequate.

Bandwidth is consumed and higher latency is avoided in this method. The issues in cloud computing can be solved by ranging it to network's edge using fog computing. Thus, it enables the services and applications of the new type. Fog computing carried extra challenges comparing to cloud computing [82]. The challenges of IoT environment and system make issues to applications in stability, total latency and reliability etc. Change in deployment can be done for the improvement of performance during runtime. Edge devices can able to connect with cloud servers anytime and anywhere. If the device cannot contain enough resources, Nowadays, IoT is considered as an integrated network framework consisting of numerous accessible devices and cloud. Edge node contains different processing capabilities [3]. IoT has altered environmental systems and operational health with devices and sensors to turn a right stream of information into location-based intelligence. With the growth of IoT requests, providing the benefits and potential of IoT technology in environment and health services is improving to develop the quality of service using devices and sensors. The network topology is very important to create a reliable network which denotes to the preparation of various requirements of IoT network by selecting the communication protocols, location & numbers of the sensors, radio range and packet size [84]. Resources rich cloud computing provides cost reduction and Quality of Service (QoS) in IoT environments. This causes in storage, analysis and processing be closer to the data formation locations and clients. So, the efficiency is improved in these resources. Every real-time IoT applications contains a set of facilities with various QoS components. By deploying the nodes in fog environment can provide the appropriate resources needed for these services [5]

Proposed System

The goal of this method is to provide farmers with more accurate information for crop forecasting purposes. For better precision, we record and save both real-time and historical data on temperature and humidity. We also collect and store rainfall data from the past. To ensure precise and accurate crop prediction, the project examines real-time data from a DHT-22 sensor measuring field temperature and humidity, historical data from a government website and/or Google Weather API, the farmer's chosen soil type, and rainfall history. Both supervised and unsupervised machine learning techniques are capable of achieving this goal. Learning networks are used to train datasets. The most accurate result is found by comparing the accuracy of several machine learning algorithms. Using learning networks, the dataset is trained. In order to provide the end user with the most accurate result, many machine learning algorithms are evaluated for accuracy. Along with the optimal crop, the algorithm also suggests the optimal fertilizer to use. A mobile-friendly, multilingual website allows farmers to interact with the system. Using a plot of land as input, the suggested algorithm will recommend the optimal crop to grow there. Examples of meteorological characteristics and soil content include precipitation, temperature, humidity, and pH. Information like as temperature, humidity, and pH is gathered by the system from either farmers or sensors.The proposed methodology for real-time prediction of crop diseases integrates Internet of Things (IoT) technologies with machine learning (ML) algorithms to enable early detection and management of plant health issues. This comprehensive approach encompasses several interconnected components, each contributing to the system's overall effectiveness.

1. Sensor Deployment and Data Acquisition: The system commences with the positioning of IoT enabled sensors on the farm field. These sensors will continuously take readings of environmental conditions like temperature, humidity, moisture, pH, and light energy. Image sensors also monitor the crop foliage for early detection of disease problems. The data will be passed wirelessly to a gateway using either of the low power consumption protocols - LoRaWAN or MQTT, which are made for long-distance communication.

2. Data Preprocessing and Feature Extraction: As the data reaches the intake point, the raw data goes through preprocessing responsible for dealing with noise, missing values, and inconsistencies. The next series of processes involves normalizing the data, smoothing and interpolating, to prepare it for analysis. Then feature extraction approaches identify the relevant characteristics essential for predicting distinct diseases, such as colour indices from images or some threshold environmental parameters.

3. Machine Learning Model Training: The preprocessed data becomes the input to a range of ML models—including Support Vector Machines (SVMs), Decision Trees, and Convolutional Neural Networks (CNNs)—that learn to classify diseases or predict their potential presence based upon patterns in the sensor data. For example, CNNs are very useful for image data as they can analyze the image data to recognize visual symptoms on leaves for detection of disease[6].

4. Real-Time Disease Prediction: Once trained, the ML models are deployed on edge devices or cloud platforms to facilitate real-time disease prediction. As new data is collected by the sensors, it is immediately processed and analyzed by the models to detect any anomalies or signs of disease. If a potential issue is identified, alerts are generated and sent to farmers or agricultural managers through mobile applications or web dashboards, enabling prompt intervention[25].

5. System Evaluation and Optimization: Metrics like accuracy, precision, recall, and F1-score are used continuously to assess the performance of the disease prediction system and all feedback through those assessments is being used for the continuous optimization of the system (e.g. retraining models with new data, changing the sensor, changing preprocessing techniques, etc.). This optimization is an iterative process to ensure the system is still working as intended and relevant to the environmental conditions and crop types for which it was devised. Integrating IoT applications with advanced ML algorithms is providing a forward-thinking approach to crop disease management that limits manual inspection of crops and provides the opportunity to course-correct when there are issues to safeguard the health and yield of our crops.

A) Digital Temperature and Humidity Sensor: It is suggested to use the DHT22 sensor to track humidity and temperature in real-time. Its accuracy and reliability have been confirmed. The data pin of the ESP32 port receives a signal from this sensor, which measures the relative humidity of the air using a thermistor and capacitive humidity sensor. The DHT22 can accurately measure humidity levels ranging from 0 to 100% RH and temperatures between 40 and 80 degrees Celsius.

B) LDR: One kind of component, known as a light-dependent resistor (LDR), may modify its resistance value in response to variations in light intensity. This allows their incorporation into circuits that detect light levels. Photoresistors (LDR) are a kind of light-dependent resistor. They have an exceptionally high thermal resistance since they are made of a semiconductor material. Light energy is transferred to electrons when photons reach the device[4]. After that, they spring into action, triggering the conductive band to become an electrical conductor.

C) Rain Sensor: A rain sensor, a kind of switching device, can detect when it is raining. Assuming it shuts when it rains, it functions similarly to a switch.

D) Rain Sensor Module: Essentially, this board works on the principle of resistance and has nickel-coated lines. With its analogue output pins, this sensor module can monitor moisture levels; when such levels are surpassed, a digital signal is outputted.

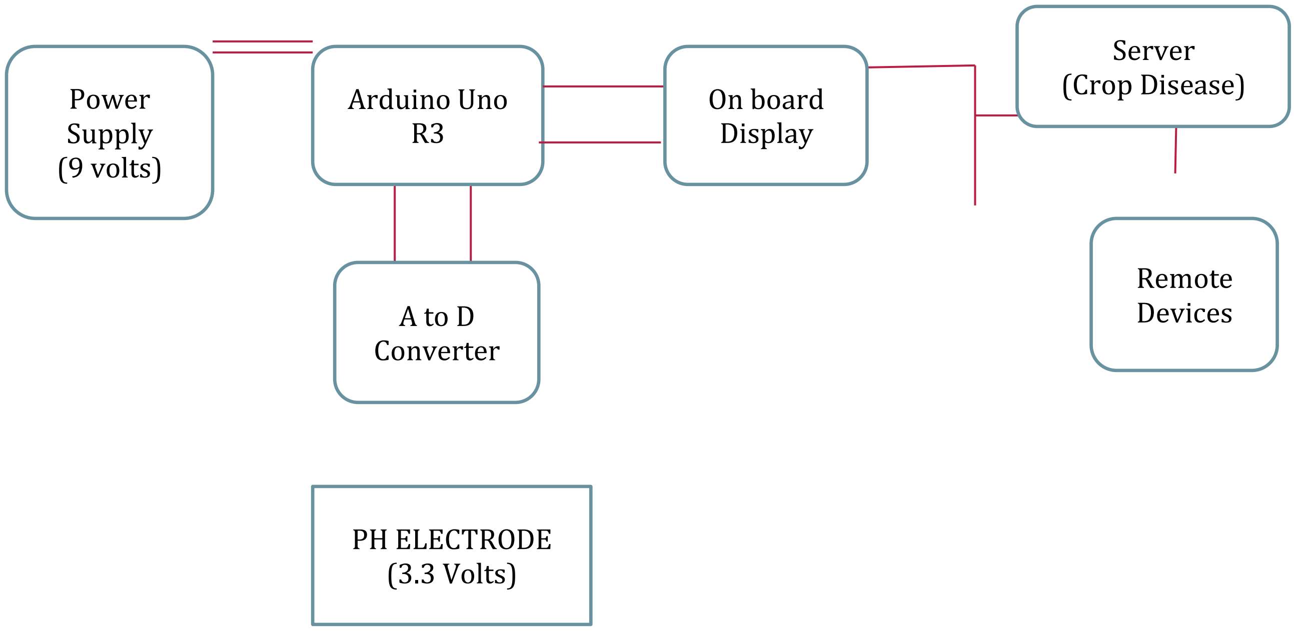

Figure 1: Block diagram of Soil Quality Analyzer and Crop Prediction System

Figure 1 shows an online crop disease monitoring system employing an Arduino Uno R3 microcontroller. The system is powered by a 9V power supply which powers the Uno R3 microcontroller. The system utilizes a 3.3V powered pH electrode sensor to measure the pH of the soil; pH is a crucial criterion to determine the overall health of the crop and potentially also if the crop is diseased. The pH electrode sends an analog signal to an Analog-to-Digital (A to D) converter, which converts the analog signal into a digital signal that can communicate with the Arduino. The Arduino can then send the digital data to an onboard display; allow local monitoring and have the ability to send digital data to a remote web server designed to collect crop disease monitoring information to be logged and used for a centralized analysis purposes. The system may also send information to remote monitors or control devices[3]. This configuration allows real-time soil condition monitoring remotely with the ability to automate data collection which creates real-time monitoring of the pH level of the soil and detection of crop conditions which have the potential to indicate disease in crop.

Working

ThingSpeak is an open-source IoT platform that enables real-time data acquisition, processing, and visualization from sensors deployed in agricultural fields. In crop disease prediction, sensors collect environmental data such as temperature, humidity, soil moisture, and leaf wetness. This data is transmitted to the ThingSpeak cloud using Wi-Fi or GSM modules. ThingSpeak channels are configured to store and organize this sensor data. The platform supports integration with Juypter Software, enabling real-time preprocessing and analysis of the incoming data. Machine learning models, either pre-trained or deployed on the cloud, are used to identify patterns linked to specific crop diseases. Thing Speak continuously monitors the sensor feeds and triggers alerts if disease conditions are detected. Its real-time graphs help farmers visualize trends and make informed decisions. The system supports scheduled data logging and automatic updates. Remote access via web or mobile apps enables farmers and agronomists to monitor field health from anywhere. ThingSpeak's open API also allows integration with third-party tools. Overall, it creates a smart, data-driven environment for proactive crop disease management.

Figure2. Different Sensor Connection for various parameter

An automated method for detecting atmospheric conditions is shown in Figure 2. Various sensors—including those for temperature, humidity, and moisture—coupled to the Arduino microcontroller. All sorts of soil and environmental parameters can be picked up by these sensors. The LCD that indicates the system's humidity, temperature, and moisture readings is shown in Figure 3.

Figure 3: LCD Readings on System

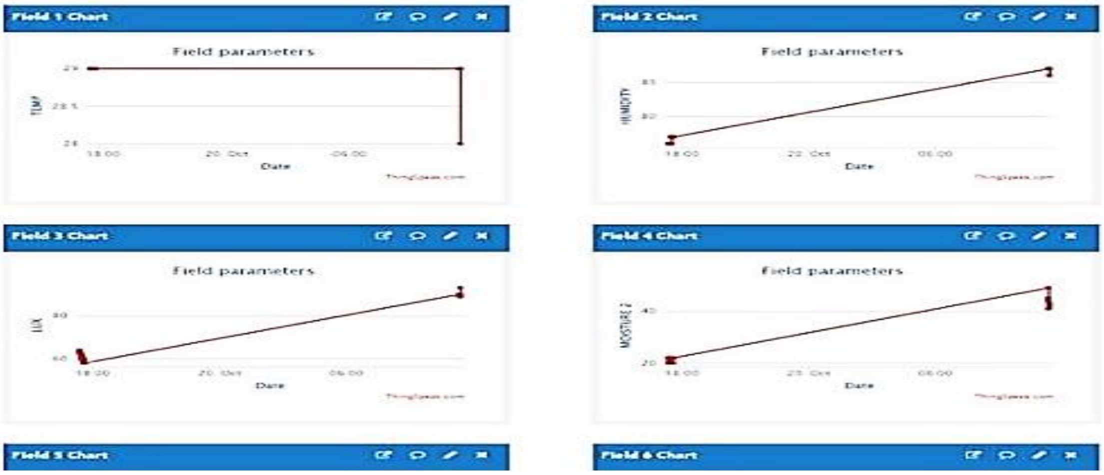

Figure 4. Real-Time Monitoring of Environmental Field Parameters Using ThingSpeak IoT Platform

Figure 4 shows the various ThingSpeak maps. Field 1, the first graph, shows the correlation between the current date and the temperature. The second graph, an illustration of the correlation between humidity and date, is the Field 2 Chart. Displayed graphically in the Field 3 Chart is the link between date and light, or lux. Field 4 Chart, the fourth graph, shows the link between date and moisture graphically.

```
Epoch 1/10
2197/2197 ━━━━━━━━━━━━━━━━━━━━ 1576s 714ms/step - accuracy:
0.3870 - loss: 2.1616 - val_accuracy: 0.8504 - val_loss: 0.4835
Epoch 2/10
2197/2197 ━━━━━━━━━━━━━━━━━━━━ 3515s 2s/step - accuracy: 0.8
360 - loss: 0.5233 - val_accuracy: 0.9100 - val_loss: 0.2758
Epoch 3/10
2197/2197 ━━━━━━━━━━━━━━━━━━━━ 1258s 572ms/step - accuracy:
0.9054 - loss: 0.2930 - val_accuracy: 0.9162 - val_loss: 0.2617
Epoch 4/10
2197/2197 ━━━━━━━━━━━━━━━━━━━━ 1245s 567ms/step - accuracy:
0.9370 - loss: 0.1939 - val_accuracy: 0.9361 - val_loss: 0.2042
Epoch 5/10
2197/2197 ━━━━━━━━━━━━━━━━━━━━ 1252s 570ms/step - accuracy:
0.9530 - loss: 0.1458 - val_accuracy: 0.9430 - val_loss: 0.1818
Epoch 6/10
2197/2197 ━━━━━━━━━━━━━━━━━━━━ 15473s 7s/step - accuracy: 0.
9625 - loss: 0.1128 - val_accuracy: 0.9492 - val_loss: 0.1811
Epoch 7/10
2197/2197 ━━━━━━━━━━━━━━━━━━━━ 5417s 2s/step - accuracy: 0.9
677 - loss: 0.0994 - val_accuracy: 0.9586 - val_loss: 0.1432
Epoch 8/10
2197/2197 ━━━━━━━━━━━━━━━━━━━━ 1333s 607ms/step - accuracy:
0.9736 - loss: 0.0791 - val_accuracy: 0.9478 - val_loss: 0.1875
Epoch 9/10
2197/2197 ━━━━━━━━━━━━━━━━━━━━ 1529s 696ms/step - accuracy:
0.9793 - loss: 0.0636 - val_accuracy: 0.9505 - val_loss: 0.1662
Epoch 10/10
2197/2197 ━━━━━━━━━━━━━━━━━━━━ 1541s 701ms/step - accuracy:
0.9812 - loss: 0.0589 - val_accuracy: 0.9656 - val_loss: 0.1210
```

Figure 5. Training Progress of Deep Learning Model Over 10 Epochs

This figure shows the training log output from a deep learning model trained using a framework like Tensor Flow or Keras over **10 epochs**. It displays key performance metrics for each epoch, which are:

- **Accuracy**: Accuracy on the training dataset.
- **Loss**: The training loss, which the model tries to minimize.
- **val_accuracy**: Validation accuracy, indicating how well the model performs on unseen data.
- **val_loss**: Validation loss, another measure of model generalization.

Final Model Performance (Epoch 10)
Training Accuracy: 98.12%
Validation Accuracy: 96.56%
Training Loss: 0.0589
Validation Loss: 0.1210
The model shows excellent generalization, with minimal overfitting and consistent improvement across epochs.

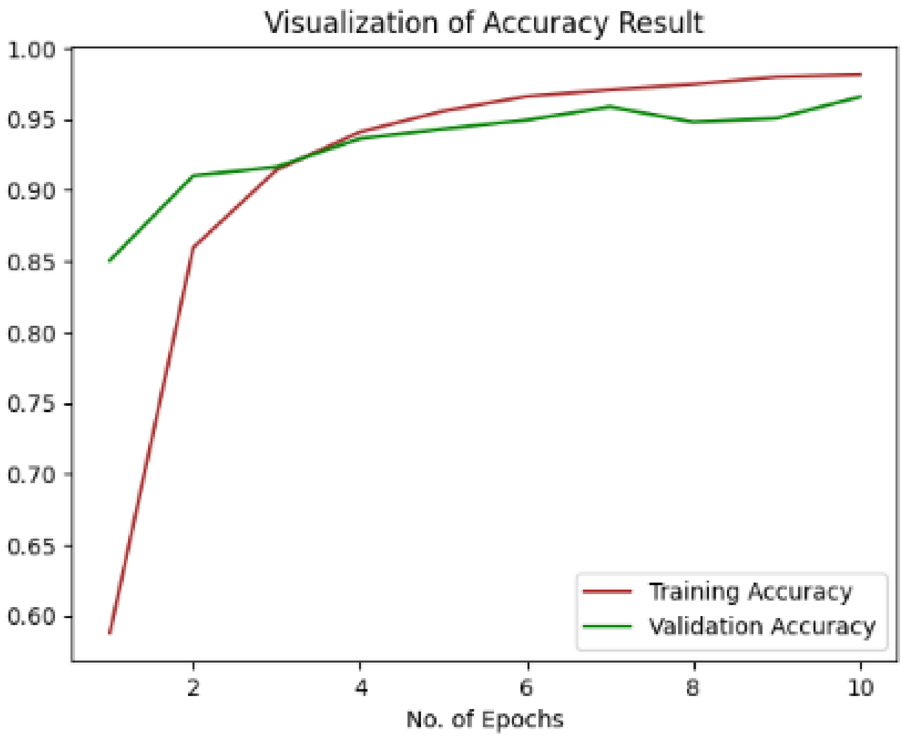

Figure 6. Accuracy Comparison Between Training and Validation Over Epochs

The "Visualization of Accuracy Result" figure 6 shows how well a deep learning model did throughout the course of ten training epochs. The training accuracy is shown in red on a line graph, whereas the validation accuracy is shown in green. From one to ten, the x-axis indicates the number of epochs. Accuracy scores ranging from 0.60 to 1.00 are shown on the y-axis.The training accuracy is rather low at epoch 1, about 0.59, which shows that the model is only starting to learn. At the same time, there is promising early generalizability, since the validation accuracy is already really high at around 0.85. Training accuracy rapidly increases across the epochs, surpassing 0.90 by the third epoch. By the third epoch, validation accuracy has also surpassed 0.91.By epoch 4, there is strong agreement, with both accuracies hovering around 0.94. The training accuracy keeps to up gradually from epoch 5 and beyond. By the tenth epoch, the training accuracy is around 0.98, which means the model is doing very well on the training data. Epoch 10 sees a general improvement in validation accuracy as well, reaching about 0.965..However, between **epochs 6 to 9**, the validation accuracy slightly fluctuates. It dips a bit around epoch 8, which may indicate a minor over fitting tendency. Despite that, validation accuracy remains high and stable, showing good generalization. The gap between training and validation accuracy is small, suggesting minimal overfitting.The consistent rise in both accuracies indicates that the model is learning well. Training accuracy always stays slightly above validation accuracy, which is expected. The close match between the two curves indicates robust model behavior. A slight drop in validation accuracy during mid-epochs is normal and not alarming. The graph confirms that the model benefits from each epoch of training. It shows that over 10 epochs, the model became significantly more accurate. The steady growth reflects a well-tuned training process. This graph is useful for evaluating how well the model fits and generalizes. Overall, the diagram demonstrates a successful and stable training phase.

	precision	recall	f1-score	support
Apple___Apple_scab	0.99	0.93	0.96	504
Apple___Black_rot	0.99	0.99	0.99	497
Apple___Cedar_apple_rust	0.98	0.99	0.98	440
Apple___healthy	0.95	0.93	0.94	502
Blueberry___healthy	0.97	0.99	0.98	454
Cherry_(including_sour)___Powdery_mildew	0.99	0.96	0.97	421
Cherry_(including_sour)___healthy	0.98	0.99	0.98	456
Corn_(maize)___Cercospora_leaf_spot Gray_leaf_spot	0.97	0.86	0.91	410
Corn_(maize)___Common_rust_	0.97	0.99	0.98	477
Corn_(maize)___Northern_Leaf_Blight	0.90	0.99	0.94	477
Corn_(maize)___healthy	0.99	1.00	1.00	465
Grape___Black_rot	0.99	0.94	0.97	472
Grape___Esca_(Black_Measles)	0.98	0.99	0.98	480
Grape___Leaf_blight_(Isariopsis_Leaf_Spot)	0.97	1.00	0.99	430
Grape___healthy	0.99	1.00	1.00	423
Orange___Haunglongbing_(Citrus_greening)	0.97	1.00	0.98	503
Peach___Bacterial_spot	0.99	0.96	0.98	459
Peach___healthy	0.99	0.99	0.99	432
Pepper,_bell___Bacterial_spot	0.98	0.94	0.96	478
Pepper,_bell___healthy	0.97	0.96	0.96	497
Potato___Early_blight	0.97	0.98	0.98	485
Potato___Late_blight	0.91	0.98	0.94	485
Potato___healthy	0.96	0.99	0.97	456
Raspberry___healthy	1.00	0.98	0.99	445
Soybean___healthy	0.93	0.99	0.96	505
Squash___Powdery_mildew	0.96	1.00	0.98	434
Strawberry___Leaf_scorch	1.00	0.91	0.95	444
Strawberry___healthy	0.99	1.00	1.00	456
Tomato___Bacterial_spot	0.98	0.94	0.96	425
Tomato___Early_blight	0.94	0.87	0.90	480
Tomato___Late_blight	0.93	0.92	0.92	463
Tomato___Leaf_Mold	0.98	0.95	0.97	470
Tomato___Septoria_leaf_spot	0.88	0.92	0.90	436
Tomato___Spider_mites Two-spotted_spider_mite	0.95	0.96	0.95	435
Tomato___Target_Spot	0.88	0.96	0.92	457
Tomato___Tomato_Yellow_Leaf_Curl_Virus	0.99	0.99	0.99	490
Tomato___Tomato_mosaic_virus	0.99	0.99	0.99	448
Tomato___healthy	0.99	0.98	0.99	481

Figure 7. Classification Report for Plant Disease Detection Model across Multiple Crop Classes

The figure presents a **classification report** that evaluates the performance of a deep learning model for plant disease detection. It includes metrics for various crop-disease class labels such as **Apple__Apple_scab**, **Tomato__Leaf_Mold**, and **Soybean__healthy**.

The evaluation metrics shown are **Precision**, **Recall**, **F1-Score**, and **Support** for each class.

From the figure, it is evident that most classes have very **high precision and recall values**, generally above 0.90.For example, **Grape__healthy** has a precision and recall of 0.99 and 1.00, indicating excellent model performance. The class **Soybean__healthy** (highlighted in the image) has a precision of 0.93, recall of 0.99, and F1-score of 0.96 across 505 samples. This indicates a slightly lower precision but very high recall for Soybean, meaning the model rarely misses healthy soybean images. Some classes, like **Tomato__Early blight**, have slightly lower recall (0.87), suggesting occasional misclassification. Meanwhile, classes like **Strawberry__healthy** and **Raspberry__healthy** show near-perfect scores. This implies the model is especially effective at detecting healthy leaves across many crop types. Overall, F1-scores are consistently high, suggesting a well-balanced model performance. High **support values** (ranging around 400–500) also show the model was tested on a large, diverse dataset. The balanced performance across both diseased and healthy categories confirms the model's generalization strength. A few slight dips in F1-score reflect challenges with visually similar diseases (e.g., different blight types).However, the overall metrics indicate that the model performs reliably across a wide range of plant species and diseases.

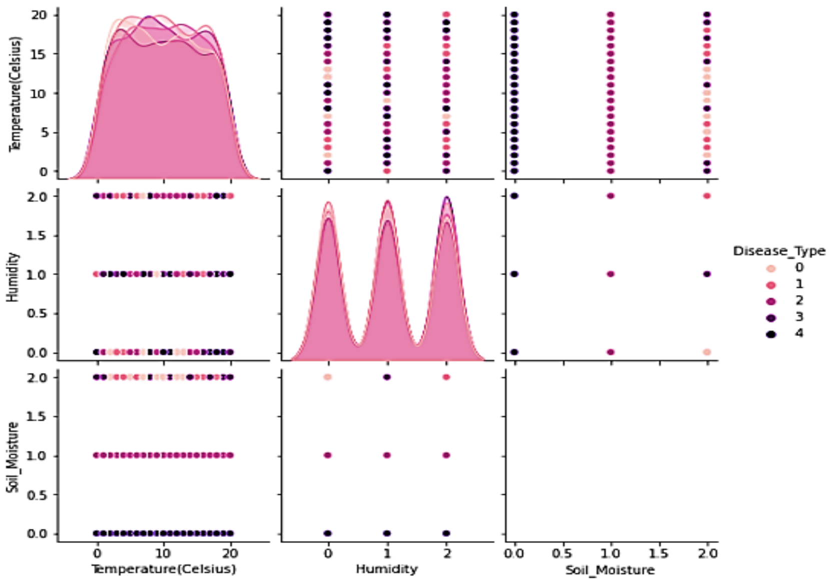

Figure 8. Pair Plot of Environmental Features Colored by Disease Type exploratory data analysis (EDA).

The figure 8 is a **pair plot (also called a scatterplot matrix)**, which visually represents the relationships between multiple numerical variables, categorized by **"Disease_Type"** using color coding. The diagonal subplots display **KDE (Kernel Density Estimation) plots** for individual variables, showing the **distribution** of each feature. For instance, the temperature has a smooth bell-shaped distribution, while humidity shows multiple peaks (multimodal distribution).

Disease patterns vary with environmental conditions: Disease types are distinguishable by different clusters or spreads in some plots. Certain diseases may correlate with specific ranges of temperature, humidity, or soil moisture.

Temperature vs. Humidity and **Humidity vs. Soil_Moisture** plots show overlapping points but reveal some structure.

Soil Moisture appears to be discretized (values at 0, 1, 2), suggesting categorical or binned data.

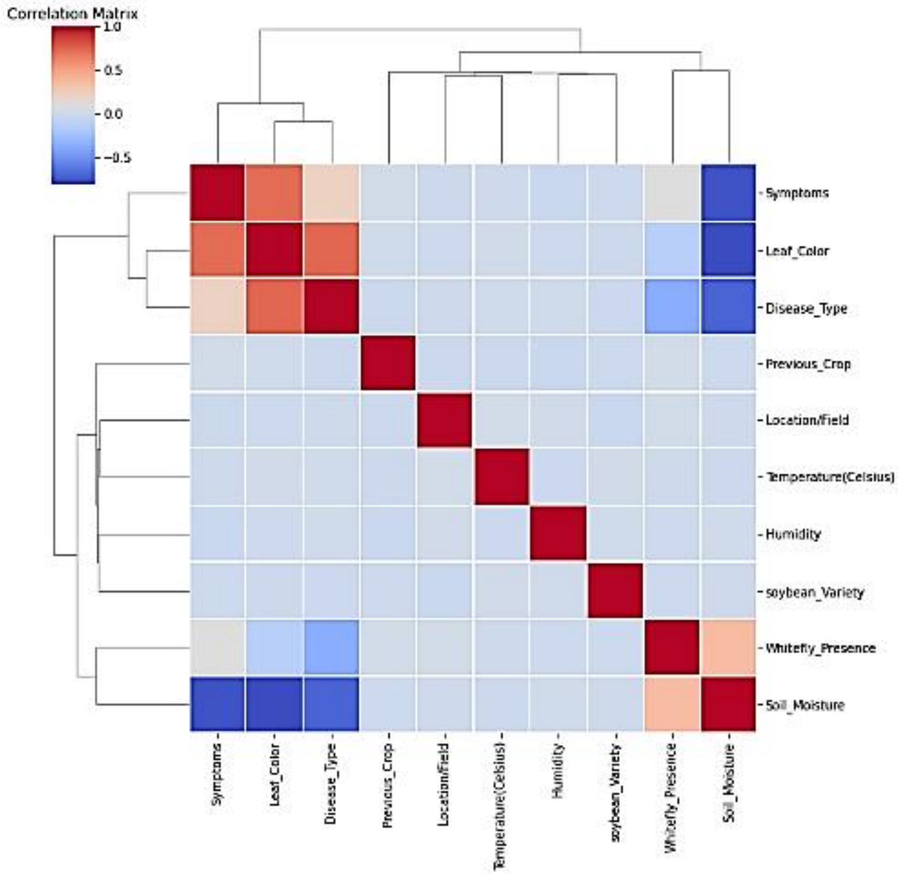

Figure 9. Hierarchical Clustered Heatmap of Correlation Matrix

Figure 9 shows a hierarchical clustering correlation heatmap that shows the direction and intensity of correlations between several environmental and agricultural factors. The correlation coefficient, which ranges from -1 (strongly negative) to +1 (strongly positive), is shown in each square of the heat map.

The color gradient indicates the strength of correlation: dark red signifies strong positive correlation, while dark blue shows strong negative correlation. The diagonal from top-left to bottom-right is deep red, indicating perfect self-correlation. Variables like Symptoms, Leaf Color, and Disease Type exhibit strong positive correlations with each other, suggesting they are closely related in disease identification. The Soil Moisture and Whitefly Presence pair also show a strong positive correlation, indicating higher whitefly presence with increased soil moisture. Environmental factors such as Temperature(Celsius) and Humidity show minimal correlation with the disease-related features, implying their impact may be indirect. The dendrograms along the top and side cluster similar variables based on correlation similarity. This clustering groups symptom-related variables together, while environmental variables form a separate group. Previous_Crop, Location/Field, and Soybean_Variety show low correlations with other features, suggesting they may be independent contributors. The mild negative correlations (in light blue) between disease indicators and environmental parameters hint at possible inverse relationships.Overall, the matrix helps identify which features are most interrelated and how they cluster, supporting decisions for feature selection in predictive modeling. The layout provides a compact and interpretable summary of complex relationships in agricultural data.

	precision	recall	f1-score	support
0	0.99	0.97	0.98	158
1	0.97	0.99	0.98	162
micro avg	0.98	0.98	0.98	320
macro avg	0.98	0.98	0.98	320
weighted avg	0.98	0.98	0.98	320

Figure 8.Binary Classification Performance Report

A binary classification model's performance summary is shown in the image. Class 0 and Class 1 both had relatively identical numbers of samples, with 158 and 162 people, respectively.

With a precision of 0.99 for Class 0, the model was able to make the majority of its positive predictions for that class come true. With a recall of 0.97, it successfully detected 97% of real class 0 instances. Class 0 has an F1-score of 0.98, indicating that recall and accuracy are well-balanced. With a recall of 0.99 and an accuracy of 0.97 for Class 1, the detection of true positives is marginally favored. Excellent model performance for both classes is confirmed by the fact that the F1-score for Class 1 is 0.98 as well. There were a total of 320 cases tested on the model; 158 were from class 0 and 162 were from class 1, as shown in the support column.

CONCLUSION

Integrating IoT-enabled data acquisition technologies with ML Models can effectively provide real-time predictions of crop diseases. Accurate environmental conditions (temperature, humidity, light intensity, and soil moisture) monitoring is vital and provides timely and accurate information for disease diagnosis. The automated IoT system continuously supports accurate data to obtain predictive variables and develop the ML model for disease diagnosis. Once trained with relevant information, ML algorithms can provide predictions of early symptoms of diseases quite accurately. This information allows farmers to take action promptly to treat the diseases and minimize crop losses. The automated system removes the human observation reliance and reduces crop losses as it improves crop quality, thereby improving farmer yield. The general visualization of the classification reports and accuracy graphs illustrates the models were correct over 80% of the time, and precision and recall remained fairly consistent across each of the crops. The Thing Speak platform was also useful in providing effective and real-time visualization and communication of the acquired data. In summary, the system is scalable, affordable, applicable across a variety of agricultural environments, and positions farmers to employ real-time, machine agnostic agriculture with improved sustainable agricultural decision making capabilities. Smart sensing with intelligent analytics was an innovation toward precision agriculture. Future work is needed to work toward developing engine models, edge computing, and expanding crop disease datasets. However, overall, improvements in sensor range and battery life also effectively improve system deployment. Lastly, the research makes a significant contribution to food security and agricultural resilience by utilizing smart technology..

REFERENCES

1. Korade, N. B., Salunke, M. B., Bhosle, A. A., Sarwade, J. M., Asalkar, G. G., Joshi, D. M., Sakure, K. S., & Sangve, S. M. (2025). Proactive Soybean Disease Detection through YOLO Leaf Extraction and ResNet-50 Classification to Reduce Crop Loss and Boost Productivity. *International Journal of Engineering Trends and Technology*, *73*(1), 385–396. https://doi.org/10.14445/22315381/ijett-v73i1p133
2. Kim, B., & Seo, S. (2025). Soybean Leaf Disease Identification Through Smart Detection using Machine Learning-convolutional Neural Network Model. Legume Research - an International Journal, Of. https://doi.org/10.18805/lrf-801

3. 3.Saini, A., Gill, N.S., Gulia, P. et al. Smart crop disease monitoring system in IoT using optimization enabled deep residual network. Sci Rep 15, 1456 (2025). https://doi.org/10.1038/s41598-025-85486-1
4. 4.R. Maurya, L. Rajput and S. Mahapatra, "RAI-Net: Tomato Plant Disease Classification Using Residual-Attention-Inception Network," in IEEE Access, vol. 13, pp. 64832-64840, 2025, doi: 10.1109/ACCESS.2025.3559804
5. 5.Shoaib M, Sadeghi-Niaraki A, Ali F, Hussain I and Khalid S (2025) Leveraging deep learning for plant disease and pest detection: a comprehensive review and future directions. Front. Plant Sci. 16:1538163.doi: 10.3389/fpls.2025.1538163
6. 6.Zainab, Z., Ali, I. a. J. A., H, B. M. M. ., Mustafa, G. K., & Jaleel, R. A. (2025). An Intelligent Model to combat Soybean Plant Disease based on Random Forest and Support Vector Machine Algorithms. Fusion Practice and Applications, 18(2), 55–65. https://doi.org/10.54216/fpa.180205
7. Santos, L. B., Gentry, D., Tryforos, A., Fultz, L., Beasley, J., & Gentimis, T. (2024). Soybean yield prediction using machine learning algorithms under a cover crop management system. Smart Agricultural Technology, 8, 100442. https://doi.org/10.1016/j.atech.2024.100442
8. 8.G, M., K S, S.K. and V A, B. (2024) 'An IOT-enabled real-time crop prediction system using soil fertility analysis', Eng, 5(4), pp. 2496–2510. doi:10.3390/eng5040130.
9. Babu GR, Gokuldhev M, Brahmanandam PS. Integrating IoT for Soil Monitoring and Hybrid Machine Learning in Predicting Tomato Crop Disease in a Typical South India Station. Sensors (Basel). 2024 Sep 24;24(19):6177. doi: 10.3390/s24196177. PMID: 39409219; PMCID: PMC11479041
10. Delfani, P., Thuraga, V., Banerjee, B. et al. Integrative approaches in modern agriculture: IoT, ML and AI for disease forecasting amidst climate change. Precision Agric 25, 2589–2613 (2024). https://doi.org/10.1007/s11119-024-10164-7
11. Kuradusenge, M., Hitimana, E., Mtonga, K. et al. SMART-CYPS: an intelligent internet of things and machine learning powered crop yield prediction system for food security. Discov Internet Things 4, 23 (2024). https://doi.org/10.1007/s43926-024-00079-0
12. Dolatabadian, A., Neik, T. X., Danilevicz, M. F., Upadhyaya, S. R., Batley, J., & Edwards, D. (2024). Image-based crop disease detection using machine learning. Plant Pathology. https://doi.org/10.1111/ppa.14006
13. 13.T. Thilagavathi, L. Arockiam, I. Priya Stella Mary, (2024). Smart Agriculture: A Comprehensive Survey on IoT-Enabled Plant Disease Detection and Agricultural Automation, 3(2), 39-47.
14. Demilie, W.B. Plant disease detection and classification techniques: a comparative study of the performances. J Big Data 11, 5 (2024). https://doi.org/10.1186/s40537-023-00863-9
15. Jafar, A., Bibi, N., Naqvi, R. A., Sadeghi-Niaraki, A., & Jeong, D. (2024b). Revolutionizing agriculture with artificial intelligence: plant disease detection methods, applications, and their limitations. Frontiers in Plant Science, 15. https://doi.org/10.3389/fpls.2024.1356260
16. Lee, S., Yun, C.M. A deep learning model for predicting risks of crop pests and diseases from sequential environmental data. Plant Methods 19, 145 (2023). https://doi.org/10.1186/s13007-023-01122-x
17. F. Qaisar, H. Shahab, M. Iqbal, H.M. Sargana, M. Aqeel, M.A. Qayyum Recent trends in Cloud computing and IoT platforms for IT management and development: a reviewPak. J. Eng. Technol., 6 (1) (Mar. 2023), pp. 98-105, 10.51846/vol6iss1pp98-105
18. Mat, I., MohdKassim, M. R., Harun, A. N., & Mat Yusoff, I. (2016). IoT in Precision Agriculture applications using Wireless Moisture Sensor Network. 2016 IEEE Conference on Open Systems (ICOS). https://doi:10.1109/icos.2016.788198
19. SachinKhirade, S. D., &Patil, A. B. (2015). Plant Disease Detection Using Image Processing. 2015 International Conference on Computing Communication Control and Automation. https://doi:10.1109/iccubea.2015.15
20. H. Al-Hiary, S. Bani-Ahmad, M. Braik and Z. ALRahamneh. Fast and Accurate Detection and Classification of Plant Diseases. International Journal of Computer Applications (0975-8887) volume 17-No.1, March 2011
21. Dr. G.H. Agrawal, Prof. S.G. Galande, Shalaka R. Londhe. Leaf Disease Detection and Climatic Parameter Monitoring of Plants Using IoT. International Journal of Innovative Research in Science, Engineering and Technology. Vol. 4, Issue 10, October 2015
22. Chougule, A., Jha, V. K., &Mukhopadhyay, D. (2016). Using IoT for integrated pest management. 2016 International Conference on Internet of Things and Applications (IOTA). https://doi:10.1109/iota.2016.7562688
23. Weizheng, S.,Yachun. W., Zhanliang, C., and Hongda, W.(2008). Grading Method of Leaf Spot Disease Based on Image Processing. In Proceedings of the 2008 International Conference on Computer Science and Software Engineering – Volume 06(December 12 – 14, 2008). CSSE. IEEE Computer Society, Washington, DC, 491494.
24. Amandeep, Arshia, Paboni. "Smart Farming Using IoT". IEEE Annual Information Technology, Electronic and Mobile Communication Conference, 8 th 2007, Vancouver BC Canada
25. DusanMarkovic, UrosPesovic. "Application of IoT in Monitoring and Controlling Agricultural Production". ActaAgriculturaeSerbica, Vol. XX, 40(2015) 145-153