Smart Irrigation Monitoring System for Precision Farming

Alin Verma[a], Navami S[a], Ayush Raj[a], Adyaman Singh[a]

[a] *Vellore Institute of Technology, Tiruvalam Rd, Katpadi, Vellore, 632014, India*

Abstract

Smart irrigation solutions, utilizing the Internet of Things (IoT) and machine learning, have the potential to revolutionize crop irrigation. These solutions use machine learning algorithms to detect real-time soil moisture levels and predict future levels. By incorporating IoT technology, farmers can remotely monitor their fields for various properties such as temperature, humidity, and soil moisture levels. The smart irrigation system then adjusts the watering schedule based on soil moisture forecasts, ensuring that crops receive the appropriate amount of water at the right time. This not only conserves water but also improves crop yield and soil health. However, the accuracy of these systems heavily depends on the quality and calibration of the sensors used. Incorrectly calibrated sensors or discrepancies in the collected data may result in inaccurate watering recommendations. For instance, unsupervised learning models, including clustering algorithms, are unreliable for consistent and accurate predictions. In contrast, supervised learning models, such as the Feedforward Neural Network (FNN) used in this study, offer greater accuracy. Experiments were conducted on a dataset, and the results clearly indicated that the FNN model and the Random Forest model outperformed other models in terms of accuracy.

Keywords:
IoT, Machine Learning, Unsupervised Learning, Irrigation System, Sensors, FNN, Random Forest

1. Introduction

Precision agriculture is an agricultural management approach that focuses on monitoring, measuring, and adapting to the variability of crops and fields. Precision agriculture entails aiding farmers in the maintenance of their crops using software that employs machine learning models built on vast sensor data.As IoT devices like actuators, and microcontrollers like Arduino Uno have become more widely available over the years, extensive use of these systems has revolutionised farming practices. The concept of "smart farming" was introduced in 2011. Since then, the use of IoT devices in agriculture has rapidly grown, and it is estimated that the global smart farming market will reach 23.14 billion USD by 2025.

Smart Farming helps farmers reduce human and intuitional errors by providing actual data and recommendations based on extensive research in the field of machine learning. Factors like low deployment cost, automation of actuators, area-specific recommendations and alerts have significantly impacted the way farming is carried out. This increase in efficiency and sustainability greatly impacts food security and enhances livestock and crop management[10]. Irrigation is one of the main operations in the farming process that needs significant labour and still relies on the farmer's instinct. Irrigation can be made more consistent and efficient by automating the entire procedure and making it more data-reliant.

Due to the extensive research and development required before implementation, these intelligent agricultural technologies lag behind other fields. Wireless Sensor Networks are an integral feature of these intelligent agriculture systems for reading and processing data from wide fields. Sensors can have limitations in accuracy, especially when exposed to environmental factors such as temperature, humidity, or other external factors[12]. Inaccurate readings can impact the overall performance and reliability of a system.

Our proposed technology intends to create a more efficient automated irrigation system and facilitate the administration of expansive fields. It incorporates research conducted over the past five years (2017-2022) to make the smart agricultural system more realisable. The suggested system can sense various data directly from agricultural fields, such as soil moisture, temperature, pH Level, and heat index, as well as extract data on the weather through evapotranspiration rates to expand the dataset for firmer results. The hardware architecture follows a master-slave architecture due to the advantageous features that is gives the system although underutilized in Precision Farming. The data collected is then processed and analysed using various ML models after which a comprehensive comparison is conducted. A lot of existing systems do not incorporate the Random Forest algorithm for predictions [7] nor do they address geographic parameters [11]. We have incorporated a Feedforward Neural Network that we designed to enhance predictions native to the soil moisture levels.

The proposed system will utilize a master-slave hardware configuration, with multiple slave nodes sending their data to respective master nodes, forming multiple clusters. This will be achieved through Access Point (AP) and Station (STA) modes, where the data will be pushed onto the cloud in real-time. Compared to conventional architectures like centralized and distributed, the master-slave approach will offer better stability, flexibility, reliability, fault tolerance, and reduced communication overhead. This approach will resolve several issues, including the single point of failure in centralization and the complex cloud storage and communication process in distribution.

The suggested system has two objectives: the accuracy of data collection and the automation of field irrigation. In addition to enhancing crop productivity, the suggested approach optimises the use of water and agricultural resources. The suggested approach coincides with the United Nations' sustainability goals, which aim to create a more sustainable future.

The remaining sections of the document are arranged in the following manner: Section 2 examines the current state-of-the-art in Smart IoT Ecosystems for Smart Irrigation and Precision Farming. A comprehensive analysis of each of these systems is presented, including a discussion of the limitations associated with our proposal based on previous studies. Section 3 introduces the framework that we have developed. Section 4 provides details on the experimental setup, results, and comparisons with existing methods and frameworks. Finally, Section 5 presents concluding remarks and suggests future research directions.

2. Literature Survey

Author Name	Techniques Used	Limitation
[1]Srishti Rawal et al (2017)	.Two YL-69 soil moisture sensors detect soil moisture levels by measuring resistance. The sensors send data to an Arduino board which controls a water sprinkler motor based on a set threshold. The Arduino board also displays the sprinkler status on a webpage.	There'sno option to manually on/off the pump Its needed in case the farmer may choose to stop growing crops or risk having the crops damaged by bad weather. Thefarmer may have to remotely shut down the system in such instances.
[2]Rajkumar et all (2017)	The proposed system provides field irrigation informationin real time. The Existing irrigation system does not make the efficient use ofwater. Here, the amount of water provided is determined by the crops' actualrequirements. This automated irrigation system optimizes resources while also cutting costs. It increases irrigation and improves environmental quality. Additionally, it reduces water waste and water scarcity.	Thereisn't a water meter installed to figure out how much water is used forirrigation and how much it will cost. Additionally, a valve should be presentso that the water flow volume can be altered. Wireless sensors should be added rather than wired ones.

Author Name	Techniques Used	Limitation
[3]Liwei Geng et al(2017)	This paper presents an intelligent agricultural monitoring system that usesa ZigBee wireless sensor network (WSN) and a deep learning algorithm. The system is designed for low power consumption and low cost. The deep learning algorithm enables the system to perform recognition and reminder functions.	Zigbee has a limited range and is less secure than WiFi. It delivers low data speed and little complexity. Its high maintenance costs, absence of a comprehensive solution, and sluggish actualization.
[4]Pavithra D et al (2017)	This paper presents a low-cost and effective GSM-based embedded system for irrigation that allows remote control of water flow via mobile devices. The system includes moisture and temperature sensors, Bluetooth for remote monitoring, and smoke sensors for detecting emergencies. It aims to improve water use efficiency and reduce waste while avoiding over- or under-irrigation.	When using GSM technologies, multiple usersaccess the same bandwidth, sometimes resulting in considerable latency as moreusers join the network. GSM technologies require carriers to install repeatersto increase coverage. Bluetooth allows only short range communication between devices.

Author Name	Techniques Used	Limitation
[5]Dweepayan Mishra et al (2018)	A breadboard connects the water pump to an Arduino board, which is also connected to soil moisture sensors to measure soil moisture content. The readings are compared to a set threshold value, and if the moisture level is too high, readings are taken periodically until it falls below the threshold. When this happens, a signal is sent to the pump to open the valve for the specific area with low moisture, allowing water to flow.	Theresno option to manually on/off the pump. InArduino sketches and shields can be difficult to modify. Arduinoboards are limited in terms of Bluetooth and Wi-Fi support .
[6]R. Nageswara Rao et al (2018)	This paper presents a simple circuitry-based irrigation system using temperature and soil moisture sensors interfaced with Raspberry Pi microcontrollers. The system provides real-time information to farmers, allowing them to make informed decisions about their crops. The proposed algorithm reduces hardware complexity and provides an all-encompassing solution to irrigation issues. Tests and observations show the effectiveness of the proposed method.	Since the model is based on cloud computing principles, it is susceptible to difficulties such as vendor lock-in and security being the greatest concern. The Raspberry Pi body is exposed to dust particles that can easily enter the device and damage the main board. Future plans include implementing the machine learning algorithm to process the data and simplifying the hardware.

Author Name	Techniques Used	Limitation
[7]Shikha Prakash et al.(2018)	The study used multiple linear regression, support vector regression, and recurrent neural networks to predict soil moisture for different periods. Three online datasets were used and multiple linear regression was found to be the best technique. Performance was evaluated using MSE.	Soil moisture prediction can be inaccurate if other factors like temperature, precipitation, and evaporation are not considered. The complex and dynamic relationship between these factors makes it difficult to standardize their interaction for accurate predictions, especially for 7-day predictions.
[8]Anna Chlingaryan et al.(2018)	Using Remote Sensing to streamline spectral and spatial data, this article discusses recent developments in machine learning for crop yield estimation and nitrogen status prediction. Review of agricultural production prediction and N status estimate, including comparative analyses of ML approaches such as BRT and CPANN.	Research lacks applications of proven ML algorithms with targeted and optimized sensors for specific PA use-cases. Also, detailed analysis is needed for the dynamic mix of stationary and mobile equipment for efficient data collection. Hybrid systems combining machine learning and signal processing methods are not fully explored.

Author Name	Techniques Used	Limitation
[9]AnnaX. Wang et al.(2018)	The study predicts soybean harvests in Argentina and Brazil using CNN and RNN trained on remote sensing data. MODIS satellite imagery was used for data collection. A pre-trained model developed in one region can be transferred to another region.	The study didn't use pre-trained models from the US which has reliable data. The test set was drawn from one particular year, which may have affected the accuracy of the model. The model was trained before the harvest date and performance was better when trained with full data.
[10]Liakos,K et al.(2018)	ML models were used in agriculture research across four categories: crop management, livestock management, water management, and soil management. ANNs were the most widely used model in crop, water, and soil management, while SVM was the most popular in livestock management. A total of eight ML models were employed.	Most studies use ANN and SVM models and neglect other pre-trained models . Data analysis takes priority over ML implementation as data grows, resulting in fewer ML applications. Approaches focus on specific parameters rather than decision-making processes.

Author Name	Techniques Used	Limitation
[11]Andrew Crane-Droesch et al (2018).	Describes an approach to yield modeling that uses a variant of a deep neural network termed semi-parametric neural networks (SNN), which can simultaneously account for complex nonlinear relationships in high-dimensional datasets, as well as known parametric structure and unobserved cross-sectional heterogeneity. Datasets of corn yield from US Midwest was used and the model outperformed classical statistical methods as well as complete non parametric neural networks performed on the same data set.	The model needs improvement for colder regions and wider geographic coverage. Climate change impact on corn yield is significant in various climate models. A research gap exists in combining elements from deterministic crop models and visualizing them statistically.
[12]Yemeserach Mekonnen et al., (2019)	The review covers the use of ML algorithms in sensor data analytics for agriculture, including a case study of an IoT-based smart farm prototype. ML techniques used with WSNs are reviewed, and an Arduino-based application is studied.	No comparison provided between ML techniques using a common dataset. Automating actuators to optimize further could be considered in the case study. Data processing performed in CIoT instead of Fog-based cloud architecture.

Author Name	Techniques Used	Limitation
[13]NTantalaki et al.(2019).	A thorough review of 106 literature was done to extract relevant data to provide a comprehensive report encompassing analytical techniques employed in precision agriculture across common themes, methodology of usage of ML techniques , implications of big data and the limitations and openissues n every field of precision agriculture.	Data privacy concerns have not been addressed in the discussed papers, and mishandling of data can happen through security holes. The large size of the datasets used increases memory and computational costs, often requiring parallel and distributed computing. Establishing a constant set of attributes that guarantee consistent and satisfactory results for all techniques can be challenging.
[14]Petteri Nevavuorib et all (2019)	A 3-layered CNN model is being used and evaluated using mean absolute error. The model has two types of errors - training error, which measures its capacity to match the data, and test error, which measures its ability to generalize to new data.	The necessity for vast quantities of labelled training data, the constraints of the model design, and the model's sensitivity to different types of data. Those who train their prediction models using remote sensing image data have achieved low prediction errors (5%), which is what the approach provided here does. As a result, it is less efficient than forthcoming models. This model is tailored to the crop and weather

Author Name	Techniques Used	Limitation
[15]Laura García et al(2020)	This document is a survey of current IoT trends in precision agriculture. Big data management and analytics for irrigation optimization are two current trends. The technique used employs a 4-layered model to produce reliable results.	Till now no model can identify water quality prior to irrigation based on the proposed architecture. Not able to transmit the information through the given frequency bandwith in the soil medium .The current implementation doesn't have blockchain and AI based system thus it lacks security and optimization .
[16]Abhinav Sharma et al(2021)	The article analyzes the use of ML in agriculture, covering topics such as yield forecasting, disease identification, and species recognition. It also explores the use of computer vision to categorize crop photos. Soil factors are predicted using ML algorithms.	Challenges in universal design of prediction algorithms due to diverse geography, accuracy dependent on quality of training dataset, sensor optimisation is crucial for smart irrigation due to erroneous sensor placement in the field.

Author Name	Techniques Used	Limitation
[17]Anneketh Vijj et al(2020)	A research study implemented a monitoring system with Raspberry Pi and sensors to address crop irrigation issues, using SVM and SVR with Radial basis function kernel for classification and quantitative forecasts of soil and crop types, and irrigation needs. Data is transmitted across a wireless sensor network and stored in the cloud.	The solution proposed by this study is gravely affected by catastrophic weather conditions. The hyper-parameters and kernel type chosen affect SVR precision. For visualisation purposes, a dedicated server or network storage is required.
[18]Rohit Sharma et al(2020)	This report proposes an ML applications framework for sustainable ASCs that provides real-time analytical insights for proactive data-driven decision-making. It offers guidelines for managing ASCs for increased agricultural productivity and sustainability.	This paper emphasizes on the application of ML for ASCs and doesn't address the source or authenticity of data. However, it lacks a clear understanding of the impact of ML on ASCs and specific guidelines for deploying ML to enhance supply chain visibility. Also, there is no customer-centric framework applicable at the user level.

Author Name	Techniques Used	Limitation
[19]Bhanu K N et al(2020)	This research explores machine learning principles for IoT-based smart agriculture systems, utilizing M2M connections and cloud computing for data processing and storage. Sensor data is sent to the cloud for ML analysis and prediction, with results used to sort the data.	The solution proposed by this study is gravely affected by catastrophic weather conditions. The hyper-parameters and kernel type chosen affect SVR precision. For visualisation purposes, a dedicated server or network storage is required.
[20]Rajendra Kumar Dwivedi et al (2020)	This research surveys various outlier detection methods for data collected by wireless sensor networks, including CESVM, KNN, PCCAD, and others. The paper concludes that there is a need for new algorithms that can outperform existing ones.	Different machine learning algorithms have different strengths and weaknesses in outlier detection, depending on the specific constraints of the problem. Hierarchical clustering may be efficient but has lower accuracy and recall rates, while other algorithms like KNN, PCA, SVM, and DBSCANS can produce more accurate results but can be resource-intensive or time-consuming.

Author Name	Techniques Used	Limitation
[21]Thomas van Klompenburga et al(2020)	This research paper is a systematic literature review that examines machine learning techniques being investigated for use in agriculture. The paper identifies gaps in the research and concludes that CNN, LSTM, and DNN are the most popular deep learning algorithms for crop yield prediction.	The quality and quantity of data features have a significant impact on the choice of deep learning methods for crop yield prediction. Overfitting and underfitting issues still exist and depend on the algorithm and feature extraction process. Feature reduction is necessary for better model training without removing important attributes. The studies reviewed vary in research scope, geographical location, and crop.
[22]Dhivya Elavarasan et al(2020)	Combining deep learning and reinforcement learning approaches has produced a framework for agricultural yield prediction that can translate raw data into crop prediction values. The proposed system creates a Deep Recurrent Q-Network model, which is an RNN over Q-learning model, to predict crop yield. The accuracy of the proposed model is 93.7%.	The proposed model in the research does not consider pests, infestations, and crop damage, which may result in significantly different outcomes from the actual situation. Additionally, the importance of statistical uncertainty is highlighted, and the need to address it for accurate target prediction is emphasized. However, the proposed model was found to have issues with exploding or vanishing gradients on larger time scales.

Author Name	Techniques Used	Limitation
[23]Meena Pundir et al (2021)	This survey paper discusses various machine learning techniques used for evaluating Quality of Service (QoS) parameters from 2011 to 2021. It covers the benefits and drawbacks of decision trees, support vector machines, K-nearest neighbor, fuzzy logic, genetic algorithms, and other ML techniques with respect to QoS-related factors like performance, privacy, and security. The paper concludes that there is a research gap in the evaluation of QoS using ML approaches and suggests further research in this area.	The report emphasises thatlimited research has been conducted on heterogeneous traffic and that each layer must satisfy QoS requirements, resulting in latency. Due to therestricted research of Machine Learning in determining QoS parameters, it is necessary to design a more pragmatic way for establishing QoS in WSN using ML.
[24]Ravesa Akhter et al (2021)	The proposed model uses IoT devices and ML for precision agriculture. It collects and analyses data from WSN devices and maintains a cloud-based database. The focus is on apple crops and detecting tree illnesses in Kashmir. The system integrates technology, knowledge mining, and visualisation techniques for reliability, optimization, security, and scalability.	There are still numerous obstacles for the suggested system, such as the difficulty of deploying in a large-scale context. Implementation cost, training model, and meteorological conditions are the parameters are going to be significant hurdles. The proposed approach is still very focused on apple cultivation in Kashmir-like environments; it must be generalised for wider application.

Author Name	Techniques Used	Limitation
[25]Khaled Obaideen et al (2022)	The paper aims to achieve sustainable irrigation using intelligent irrigation systems. It proposes a wireless sensor network and IoT system for automatic irrigation, requiring minimal human intervention. A comparative analysis of technologies, including sensors and communication devices, is undertaken, and machine learning techniques such as neural networks and fuzzy logic are evaluated.	Controlling and safeguarding such massive IoT systems remains challenging. The major objective of the proposed work is sustainability; nevertheless, if the system requirements are more destructive to the environment than the advantages they provide, the entire system is declared infeasible. In an automated system, it is crucial that the results of data processing using ML models be accurate; otherwise, a crisis could develop.

The selected pieces of literature (25 publications) were studied in detail to extract the information relevant to our research issues. We focused on papers that discussed in detail various ML techniques for feature prediction, existing smart irrigation architectures and usage of WSN and IoT infrastructures for the implementation of proposed systems. In the papers analyzed, we observed that there is a lack of reliability on WSN sensors in situations persisting in environmental changes [12]. Optimization of sensors while building a smart irrigation system is crucial as the accuracy of the system can be affected due to erroneous sensor placement in the field [17]. The data set generated can be unexpectedly altered due to sudden events in weather conditions at which state the present systems may fail which demands high-sensitive sensors to be protected [16]. Random Forest Algorithm is effective in irrigation-based architecture and needs to be employed in modelling and training data sets [19]. Rajkumar et al [2] proposed irrigation system but , there wasn't a mechanism to manually handle motor pump functions. The lack of a meter for observing features like the volume of water left for irrigation and the cost of its usage is present [1]. An additional valve for altering water flow should also be present along with a secondary pump in case of primary pump failure. These limitations were concluded from the surveyed studies.

3. Proposed Work

3.1. Methodology

The proposed methods are meant to benefit the agricultural sector and help farmers lighten their workload, particularly in labour-intensive tasks like irrigation. The project aims to make agriculture more data-driven and manageable by removing the element of unpredictability. Additionally, it aims to improve the usability and interactivity of intelligent agriculture systems by automating the responses of machinery to critical conditions. To ensure enough coverage and water flow to all locations in the field, the method will deploy a reliable water distribution system. The information gathered from numerous nodes dispersed around the field equipped with sensors for temperature, humidity, wetness, etc. will be processed to determine what is causing the water level to decrease or grow so that decisions can be made. The deployment of actuators that may also be operated manually will make handling during severe weather conditions easier. Additionally, a prediction system employing machine learning methods like Random Forest and our Feedforward Neural Network framework will be applied to the data set obtained to give farmers more specific predictions like the anticipated crop output.

3.2. Hardware Design

The main hardware components of the System are –

- ESP32 NodeMCU
- DHT22 sensor
- Soil Moisture sensor
- pH sensor
- ServoMotor
- Breadboard

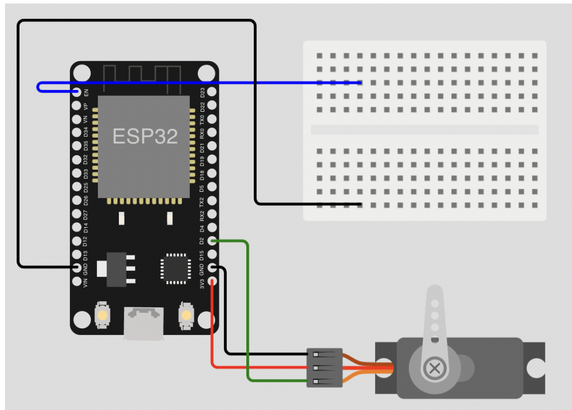

Figure 1: A server node consisting of a ESP32 WROOM 32 microcontroller and a servo motor

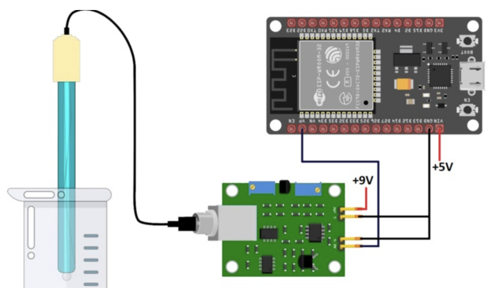

Figure 2: ESP32 Devkit V1 client node with the pH sensor setup

- Jumper Wires

We take the breadboard and create connections using it. We use 2 NodeMCUs for our setup, one that acts as a server and another as the client. The server has the Servo motor attached to it via jumper wires. We use the D2 pin of the NodeMCU for the same. The client node has all the other sensors. The DHT22 sensor is connected via the D15 pin and the moisture sensor via D2 and the pH sensor via the D13 pin. We use a signal conversion board attached to the pH electrode for the appropriate signal transmission. Capacitors have been placed to enable the download mode of Arduino IDE. A sample of soil was acquired and the sensors were put in place. The client and the server were connected to the laptop to upload the code and run the devices. The sensor network deployed is done through a server-client node architecture, where each region of the target field has several client

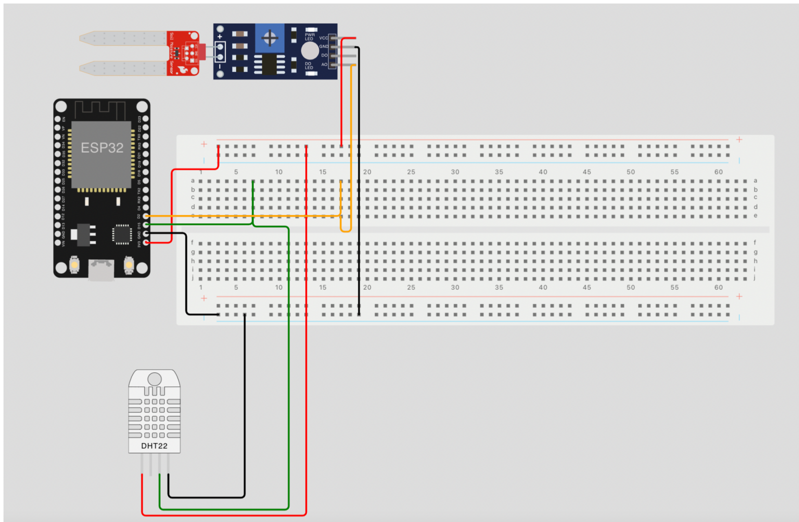

Figure 3: ESP32 Devkit V1 client node with a DHT22 sensor and a soil moisture sensor

nodes which read data and forward it to the server which uploads it to the FireBase cloud in regular time intervals, where the data is stored in NoSQL format.

In Figure 4, the sensors are reading the environment and transmitting it to the server node which will further forward it to the cloud.

3.3. Dataset and Machine Learning Model

In a soil moisture prediction task, the input layer is followed by one or more hidden layers and an output layer, where the output layer produces a continuous numerical value representing the predicted soil moisture. The attributes are:

- NPK Levels:-
    1. Nitrogen Content of Soil(N)
    2. Potassium Content of Soil(P)
    3. Phosphorus Content of Soil(K)
- The surroundings Temperature(Temperature)
- Evaporation Rate(humidity)
- Amount of Precipitation(Rainfall)

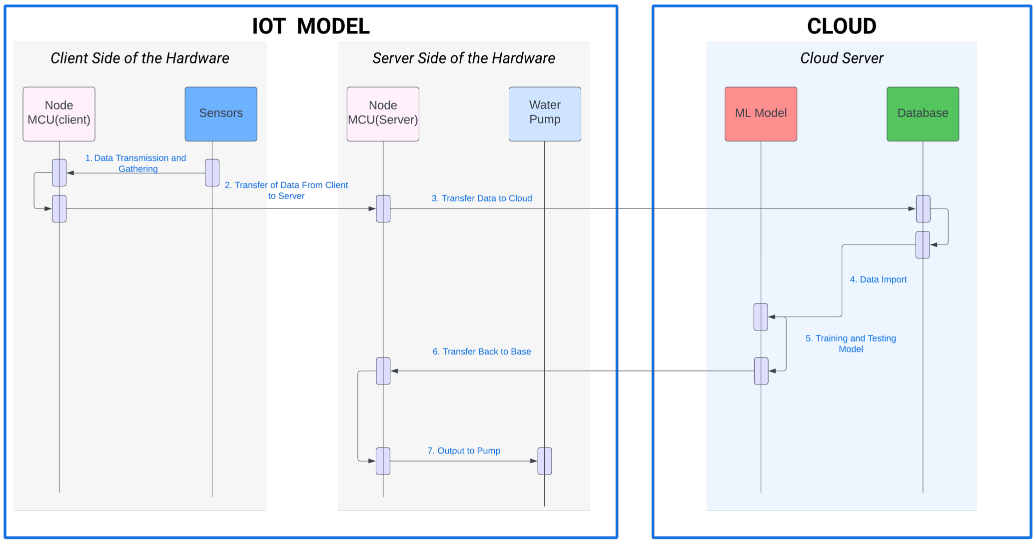

Figure 4: A sequence diagram depicting the workflow of the experimental setup

- The Crop(Label)
- Soil Moisture Level(Moisture)

Random Forest Regression is a highly effective machine learning algorithm utilized for making predictions on continuous numerical data. It works by constructing multiple decision trees and aggregating their predictions to arrive at the final output. Random Forest Regression has proven to be an accurate and reliable method for a wide range of prediction tasks, including soil moisture prediction. The algorithm can handle complex interactions between input variables and provide robust predictions even when dealing with noisy or incomplete data. By using a large number of decision trees, Random Forest Regression can capture the variability in soil moisture levels due to various factors such as temperature, precipitation, and soil type, resulting in highly accurate predictions. Overall, Random Forest Regression is an excellent tool for soil moisture prediction, providing a reliable and accurate method for researchers and farmers to monitor and manage soil moisture levels.

3.4. Integration to Cloud

The uploading of data collected by a sensor network in a farm to Firebase Realtime Database can be a crucial step in the implementation of an effective irrigation management system. Firebase Realtime Database is a cloud-based database that allows for real-time synchronization and easy data retrieval. With Firebase Realtime Database, data from the sensor network can be easily stored and accessed from anywhere, enabling the development of machine learning models for predicting soil moisture levels and making irrigation decisions.

The ML model can be trained using historical sensor data stored in the Firebase Realtime Database to learn the relationships between environmental factors such as temperature, humidity, and soil moisture, and to determine the optimal irrigation schedule for the crops. The predictions generated by the ML model can be used to control the motor responsible for irrigation, ensuring that crops receive the necessary amount of water at the right time.

By using Firebase Realtime Database for data storage and retrieval, farmers can benefit from an intelligent irrigation management system that helps improve crop yield and conserve water resources. The real-time synchronization feature of Firebase Realtime Database also allows for immediate adjustments to the irrigation schedule, based on real-time sensor data, to optimize crop growth and minimize water waste. In table 1, we have drawn contrast of proposed system to earlier proposed systems on

4. Discussion and Results

4.1. Machine Learning Models

The literature Survey conducted concluded that the most common models used for such Agriculture Soil based datasets are SVM(Support Vector Mechanism), Gradient Boosting Regression, and Random Forest Regression.

SVM: Support Vector Machine (SVM) is a highly effective machine learning algorithm used for classification and regression tasks. The SVM algorithm identifies a hyperplane that separates the input data into two classes with maximum margin. This hyperplane is chosen in a manner that maximizes the distance between the nearest points of the two classes. The SVM algorithm is especially helpful when the data is not linearly separable. In such scenarios, SVM can use a kernel function to transform the input data into a higher-dimensional space where it can be linearly separated. SVM has

Table 1: Comparison of Proposed System to Other Proposed Systems

Functions	[8]	[12]	[14]	[16]	Proposed System
Time Based/ Environment Based	Time Based	Environment Based	Time Based	Time Based	Environment Based
Focus	Soil moisture Prediction	Sensor Network Deployment	Crop Yield Prediction	Crop Yield Prediction	Soil Moisture Prediction
Dataset Attributes	Soil Moisture, Time Stamp	Temp., Yield, Soil Moisture	Yield, Variety, Imaging Date, Growth Phase	Temp., Yield, Humidity, Gas Soil Moisture	Soil Moisture, N, P, K, Temp., Humidity, Variety
ML Models Studied	Multiple Linear Regression, SVM,	SVM, Random Forest, ANN	CNN	SVM, Random Forest, Gradient Boosting	SVM, Gradient Boosting, FNN, Random Forest
Final ML Models Used	Multiple Linear Regression	-	CNN	SVM	Random Forest Regression

found successful applications in several areas, such as image classification, speech recognition, and anomaly detection. Its capability to handle complex datasets and nonlinear relationships between input variables has made it a preferred choice for many machine-learning tasks. SVM is a robust and powerful tool for machine learning that offers high accuracy in various applications

Gradient Boosting Regression: Gradient Boosting Regression is a machine learning algorithm that is widely used for regression tasks due to its popularity and effectiveness. It is based on combining multiple weak learners, such as decision trees, to form a strong learner that can accurately predict the target variable. The algorithm works by fitting decision trees sequentially, with each new tree correcting the errors of the previous ones. This approach leads to a highly accurate model that can deal with complex nonlinear relationships between the input and target variables. Gradient Boosting Regression has found successful applications in predicting housing prices, stock prices, and customer churn, among other problems. The algorithm is particularly advantageous in handling large datasets with many features, as it can manage high-dimensional input spaces and perform feature selection. Overall, Gradient Boosting Regression is a versatile and robust machine learning algorithm that can be applied to various regression tasks..

Random Forest Regression: Random Forest Regression is a well-known and potent machine learning algorithm utilized for forecasting continuous numerical values. It works by constructing multiple decision trees and aggregating their predictions to arrive at the final output. The algorithm creates each decision tree using a subset of the available input features, and it randomly selects the subset for each tree. This approach helps to reduce overfitting and improve the generalization performance of the model. Random Forest Regression is highly accurate and reliable, even when dealing with noisy or incomplete data. It is particularly useful when dealing with high-dimensional input spaces, as it can perform feature selection and identify the most important variables for predicting the target variable. Random Forest Regression has been successfully applied to a wide range of applications, including finance, healthcare, and agriculture. In general, Random Forest Regression is an adaptable and potent machine learning technique that can be employed in various prediction tasks, delivering precise and dependable predictions across multiple domains.

model	MAE	MSE	RMSE	R2
Support Vector Mechanism	0.0371402	0.00200696	0.0447991	0.999787
Gradient Boosting Regression	0.0371402	0.00200696	0.0447991	0.999787
Random Forest Regression	0.00930474	0.00087559	0.0295904	0.999907
Feedforward Neural Network	0.0308043	0.00168458	0.0410436	0.998315

Figure 5: A statistical comparison of the four models employed

Apart from the models concluded from Literature Survey, another machine learning not explored by the previous work is the use of Neural Networks in performing such regression problems. Feedforward Neural Networks is a kind of neural network that can be developed and trained to perform with higher Accuracy. Recurrent Neural Networks and Convolutional Neural Networks were specifically removed from consideration due to their proficiency in handling other kinds of data types rather than tabular data.

Feedforward Neural Networks: Feedforward Neural Networks have emerged as a useful tool in predicting soil moisture levels for agricultural purposes. The neural network structure comprises multiple layers of interconnected neurons, where each neuron in a layer receives input from the previous layer. To predict soil moisture levels, the input layer is followed by one or more hidden layers and an output layer, which generates a numerical value representing the predicted soil moisture level. During the training process, the network acquires knowledge about the correlation between environmental factors, such as temperature, humidity, and precipitation, and soil moisture levels, by adjusting the weights of the neuron connections to minimize the deviation between the predicted and observed soil moisture values. Due to their flexibility and capacity to model intricate relationships between environmental factors and soil moisture levels, Feed forward Neural Networks have the potential to predict soil moisture in various agricultural applications.

Figure 5 shows the error values of different machine learning models, from which we can conclude that Random forest Regression has shown the best results and will be further employed in our project to appropriate soil

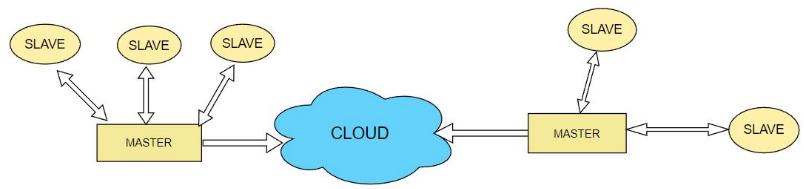

Figure 6: master-slave architecture

moisture values for different crops in different environments.

4.2. Sensor Networks

The sensors were able to detect the values accurately and precisely as shown in Figure 6. The data sent to the server was seen in our local host and the values were successfully pushed to Firebase for the machine learning models to decide in real-time. The motor on receiving the value 1, starts turning its shafts according to the conditions we provided. This functional system, in terms of hardware and software, enables us, to have a clear enclosure on the soil conditions. It is found that the system can be deemed useful for real-time analysis of the crop's condition at a low-cost. The NodeMCU was a fair cheaper option as compared to using an Arduino Uno board with external WiFi modules and this has them in-built. The client nodes can be placed far from each other making it commercially viable as well.

We are utilizing a master-slave approach for our hardware configuration, where multiple slave nodes will send their data to their respective master node, forming multiple clusters. This is done in two modes; Access Point (AP) and Station (STA). In AP mode, the slaves send their data to the master node via WiFi modules, which in STA mode, is pushed onto the Cloud in real-time in which it connects to the Internet Hotspot.Conventional architectures like centralised and distributed has several issues persistent to them which is significantly resolved in master-slave approach. Centralised has a central hub acting as the single point of failure whereas in master slave nodes operate independent of each other and functions even if master fails. In distributed, each end device collects and analyses data by itself. this results in a complex cloud storage and communication process making it computationally hard to collect data from such large sources through individual connections leading to network disruptions. Thus the master-slave approach accounts for better

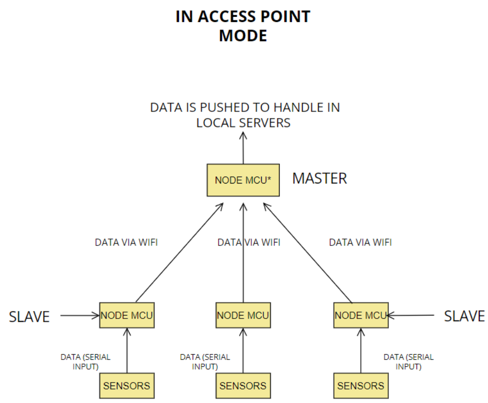

Figure 7: Access Point Mode

stability and flexibility, reliability, increased fault tolerance, reduced risk of system downtime and a significant reduction in communication overhead.

The advantages of the master-slave architecture include-

- Scalability: The master-slave architecture's scalability allows for easy addition of client nodes to the system without impacting the server's functionality, making it ideal for large-scale agricultural fields where multiple sensors may be required to cover a wide area. In centralised this poses an issue as the data transmission and storage increases greatly. In distributed, monitoring all the nodes gets difficult.

- Efficient Communication: By optimizing the server's ability to send commands and instructions to the client nodes, the need for continuous bidirectional communication and conserving resources like power and bandwidth reduces. In distributed system, the overhead gets too complex to be handled in an efficient way.

- Hybrid Control: The master-slave architecture combines the benefits of both centralized and distributed systems, where slave nodes operate independently, reducing the impact of a single failure, and communication occurs among multiple master nodes instead of a single central hub, improving data handling capabilities.

- Security: The master-slave architecture can improve system security by restricting access and control of client nodes. The server can act as a

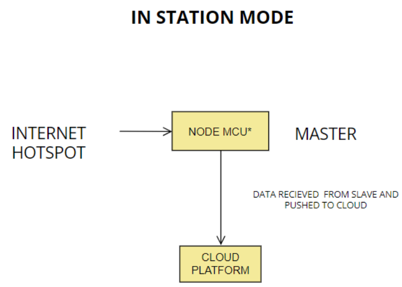

Figure 8: Station Mode

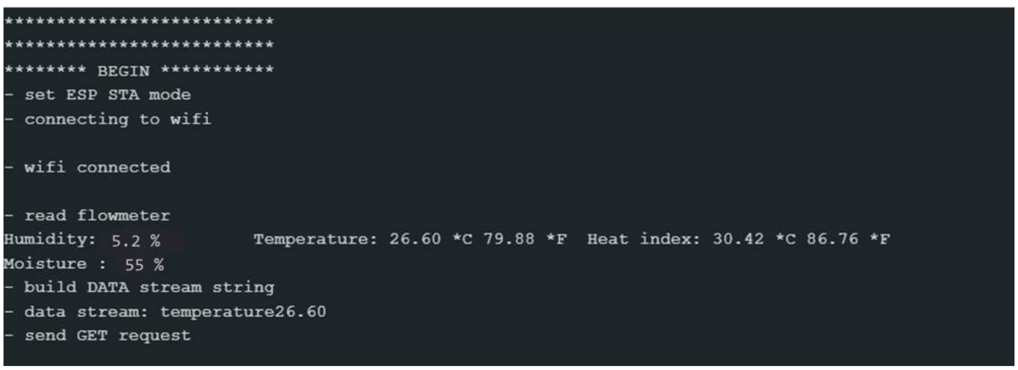

Figure 9: Result from the serial monitor of the Arduino IDE

gatekeeper, monitoring and controlling actions of client nodes, thereby preventing unauthorized access or tampering with the system.

5. Conclusion

This research paper presented a Smart Irrigation Farming System that utilized various sensors for soil moisture level detection and used the acquired data on algorithms like the Random Forest algorithm for accurate results. Moreover, the system incorporated a motor that ran when the ML model provided a positive output, which is an indication of a low water level, ensuring that the plants received an adequate water supply. The motor was triggered by the moisture sensor, which was connected to the microcontroller and controlled the water pump accordingly. This automated system provided a more efficient and accurate irrigation method than traditional manual systems, as it prevented under or over-watering of the plants, which can negatively impact crop yield. The use of a Random Forest algorithm also contributed to the system's accuracy, as it enabled the prediction of soil moisture content by analyzing the sensor readings and other environmental factors such as temperature, humidity, and sunlight. The results showed that the random forest algorithm was able to accurately predict soil moisture content with an R-squared value of 0.99. Overall, this research highlights the effectiveness of integrating advanced technologies and machine learning algorithms into traditional irrigation systems, leading to improved efficiency, reduced water consumption, and increased crop yield. These findings contribute to the development of smart agriculture practices that are essential for sustainable and efficient agriculture in the face of climate change and population growth.

6. Future Works

In future, other environmental factors such as temperature, humidity, and sunlight can also affect crop growth and yield. Future research could explore the integration of sensors to monitor these factors and develop a more comprehensive and detailed smart irrigation system. Sensors consume a lot of power for functioning and require frequent maintenance.E mploying the use of EMI sensors will widely help tackle this issue. EMI sensors measure the electrical conductivity of soil, which is related to soil moisture content. By measuring the electrical conductivity at different depths and locations, EMI sensors can provide a detailed map of soil moisture across a field. This can help to identify areas with high or low soil moisture and adjust irrigation accordingly. EMI sensors are severely underutilized in soil moisture monitoring and their various advantages can be used to leverage the current issues with data consistency. EMI sensors being non-invasive doesn't require contact with soil or water for measurement eliminating environmental sensitivity faced by other sensors used. Added to that, EMI sensors use a high-frequency electromagnetic field to measure the water level. This method is highly accurate and can detect changes in water levels as small as a few millimetres. The integration of Fog and Edge cloud computing can enable the collection and analysis of large amounts of data from multiple sensors in real time. This can lead to the development of more intelligent and responsive irrigation systems that can adapt to changing environmental conditions and is scalable and efficient.

References

[1] Rajkumar, M. Newlin, S. Abinaya, and V. Venkatesa Kumar. "Intelligent irrigation system—An IOT based approach." 2017 International Conference on Innovations in Green Energy and Healthcare Technologies (IGEHT). IEEE, 2017.

[2] Rawal, Srishti. "IOT based smart irrigation system." International Journal of Computer Applications 159.8 (2017): 7-11.

[3] Geng, Liwei, and Tingting Dong. "An Agricultural Monitoring System Based on Wireless Sensor and Depth Learning Algorithm." Int. J. Online Eng. 13.12 (2017): 127-137.

[4] Pudumalar, S., et al. "Crop recommendation system for precision agriculture." 2016 Eighth International Conference on Advanced Computing (ICoAC). IEEE, 2017.

[5] Mishra, Dweepayan, et al. "Automated irrigation system-IoT based approach." 2018 3rd International conference on internet of things: Smart Innovation and Usages (IoT-SIU). IEEE, 2018.

[6] Rao, R. Nageswara, and B. Sridhar. "IoT based smart crop-field monitoring and automation irrigation system." 2018 2nd International Conference on Inventive Systems and Control (ICISC). IEEE, 2018.

[7] Liakos, Konstantinos G., et al. "Machine learning in agriculture: A review." Sensors 18.8 (2018): 2674.

[8] Prakash, Shikha, Animesh Sharma, and Sitanshu Shekhar Sahu. "Soil moisture prediction using machine learning." 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT). IEEE, 2018.

[9] Chlingaryan, Anna, Salah Sukkarieh, and Brett Whelan. "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review." Computers and electronics in agriculture 151 (2018): 61-69.

[10] Crane-Droesch, Andrew. "Machine learning methods for crop yield prediction and climate change impact assessment in agriculture." Environmental Research Letters 13.11 (2018): 114003.

[11] Wang, Anna X., et al. "Deep transfer learning for crop yield prediction with remote sensing data." Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies. 2018.

[12] Mekonnen, Yemeserach, et al. "Machine learning techniques in wireless sensor network based precision agriculture." Journal of the Electrochemical Society 167.3 (2019): 037522.

[13] Tantalaki, Nicoleta, Stavros Souravlas, and Manos Roumeliotis. "Data-driven decision making in precision agriculture: The rise of big data in agricultural systems." Journal of Agricultural & Food Information 20.4 (2019) : 344-380.

[14] Nevavuori, Petteri, Nathaniel Narra, and Tarmo Lipping. "Crop yield prediction with deep convolutional neural networks." Computers and electronics in agriculture 163 (2019): 104859.

[15] García, Laura, et al. "IoT-based smart irrigation systems: An overview on the recent trends on sensors and IoT systems for irrigation in precision agriculture." Sensors 20.4 (2020): 1042.

[16] Vij, Anneketh, et al. "IoT and machine learning approaches for automation of farm irrigation system." Procedia Computer Science 167 (2020): 1250-1257.

[17] Sharma, Abhinav, et al. "Machine learning applications for precision agriculture: A comprehensive review." IEEE Access 9 (2020): 4843-4873.

[18] Sharma, Rohit, et al. "A systematic literature review on machine learning applications for sustainable agriculture supply chain performance." Computers & Operations Research 119 (2020): 104926.

[19] Bhanu, K. N., H. J. Jasmine, and H. S. Mahadevaswamy. "Machine learning implementation in IoT based intelligent system for agriculture." 2020 International Conference for Emerging Technology (INCET). IEEE, 2020.

[20] Dwivedi, Rajendra Kumar, Arun Kumar Rai, and Rakesh Kumar. "Outlier detection in wireless sensor networks using machine learning techniques: a survey." 2020 International Conference on Electrical and Electronics Engineering (ICE3). IEEE, 2020.

[21] Van Klompenburg, Thomas, Ayalew Kassahun, and Cagatay Catal. "Crop yield prediction using machine learning: A systematic literature review." Computers and Electronics in Agriculture 177 (2020): 105709.

[22] Elavarasan, Dhivya, and PM Durairaj Vincent. "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications." IEEE access 8 (2020): 86886-86901.

[23] Pundir, Meena, and Jasminder Kaur Sandhu. "A systematic review of quality of service in wireless sensor networks using machine learning: Recent trend and future vision." Journal of Network and Computer Applications 188 (2021): 103084.

[24] Akhter, Ravesa, and Shabir Ahmad Sofi. "Precision agriculture using IoT data analytics and machine learning." Journal of King Saud University-Computer and Information Sciences 34.8 (2022): 5602-5618.

[25] Obaideen, Khaled, et al. "An overview of smart irrigation systems using IoT." Energy Nexus (2022): 100124.