Research Article GU J Sci, Part A, 9(4): 378-391 (2022) https://doi.org/10.54287/gujsa.1141575

Gazi University
Journal of Science
PART A: ENGINEERING AND INNOVATION
http://dergipark.org.tr/gujsa

Machine Learning Techniques for the Classification of IoT-Enabled Smart Irrigation Data for Agricultural Purposes

Aamo IORLIAM[1*] Sylvester BUM[2] Iember S. AONDOAKAA[3] Iveren Blessing IORLIAM[4] Yahaya SHEHU[5]

[1,4]Centre for Food Technology and Research, BSU, Makurdi, Nigeria
[1,2,3]Department of Mathematics & Computer Science, BSU, Makurdi, Nigeria
[5]Shehu Shagari College of Education, Sokoto, Nigeria

Keywords

IoT

Agriculture

Smart Irrigation

Machine Learning Techniques

Abstract

To support farming year-round, a variety of smart IoT irrigation devices have recently been developed. It is crucial to forecast the soil moisture of agricultural farms so as to produce high yields since the high yields depends on the efficiency of water supply on farmlands. In smart irrigation, anytime water is needed on the farms, the smart pumps switch on to pump the required water so as to prevent the crops from drying up. The smart pumps also shut down if the farms have the ideal level of soil moisture, preventing over-flooding of the fields. Data is generated when the smart pumps are ON or OFF at any given time. Therefore, it is crucial to classify the data produced by smart IoT-enabled irrigation devices when these devices are ON or OFF. In this paper, the soil moisture, temperature, humidity, and time are used as inputs into machine learning techniques for classification. These machine learning techniques include logistic regression, random forest, support vector machine, and convolutional neural network. According to experimental findings, the accuracy of the logistic regression was 71.76%, that of the random forest was 99.98%, that of the support vector machine was 90.21%, and that of the convolutional neural network was 98.23. Based on the high accuracy that the random forest attained, it has more potential to help in assessing smart irrigation conditions (wet or dry) in an optimized manner.

Cite

Iorliam, A., Bum, S., Aondoakaa, S. I., Iorliam I. B, & Shehu, Y. I. (2022). Machine Learning Techniques for the Classification of IoT-Enabled Smart Irrigation Data for Agricultural Purposes. *GU J Sci, Part A*, *9*(4), 378-391.

Author ID (ORCID Number)

A. Iorliam, 0000-0001-8238-9686
S. Bum, 0000-0003-3342-6457
I. S. Aondoakaa, 0000-0002-0812-0109
I. B. Iorliam, 0000-0002-9973-6151
Y. I. Shehu, 0000-0001-8924-9344

Article Process

Submission Date	06.07.2022
Revision Date	08.08.2022
Accepted Date	07.10.2022
Published Date	31.12.2022

1. INTRODUCTION

Crop growth is now severely hampered by the relative scarcity of water brought on by drought or insufficient rainfall throughout the world (Goap et al., 2018). Irrigation is the process of delivering the necessary amount of water to plants at predetermined intervals. It is necessary for both farming and raising agricultural output and it has historically been the solution to the problem of water scarcity on farms (Surendran et al., 2015). Prior to today, farms have relied on conventional irrigation techniques including furrow irrigation, check basin irrigation, strip irrigation, and basin irrigation to lessen the impact of water scarcity. To irrigate crops, traditional irrigation techniques included buckets, pumps, canals, and watering cans. This approach frequently requires a lot of labor to convey water to the roots of the plants (Janani & Jebakumar, 2019).

Research has produced better irrigation techniques known as "smart irrigation" using the Internet of Things (IoT) to control water scarcity on farms. Utilizing sprinklers, pipes, ditches, and other similar equipment, smart irrigation distributes a controlled amount of water at the proper moment (Adam et al., 2020). This aims to increase crop yields while conserving water and money (Iorliam et al., 2021). Smart irrigation systems that use machine learning are more accurate at estimating how much water plants need (Çetin & Beyhan, 2022). In order to successfully address the issue of inadequate or excessive irrigation, these machine learning techniques

*Corresponding Author, e-mail: aamoiorliam@gmail.com

consider agricultural parameters to effectively curtail the issue of under or over irrigation (Çetin & Beyhan, 2022). Different machine learning techniques have been created or adopted by researchers in order to create smart irrigation systems that make use of wireless sensor networks, big data, cloud computing, and the internet of things in order to efficiently manage the water requirements of crops on smart farms (Sanjeevi et al., 2020).

In order to replicate human behaviors and make informed choices, decision-support tools have used irrigation data from human specialists to comprehend the scheduling patterns of irrigation systems (Pluchinotta et al., 2018). These tools continuously learn from their experiences and improve their performances. For the goal of smart irrigation management, machine learning methods such as artificial neural networks, support vector regressions, decision trees, random forests, and naive bayes have been applied (Goap et al., 2018; Romero et al., 2018; Nawandar & Satpute, 2019). By utilizing computational tools to learn from irrigation data, these machine learning algorithms or techniques play significant roles in smart irrigation systems. The main goal of this learning process is to train IoT-enabled data in order to carry out smart irrigation operations (Janani & Jebakumar, 2019). The collection of the training data is determined by some features, which may be binary, nominal, or numeric (Janani & Jebakumar, 2019). Therefore, to classify IoT-generated data from smart irrigation systems, this research uses machine learning techniques such as logistic regression (LR), random forest (RF), support vector machine (SVM), and convolutional neural network (CNN) algorithms. In order to identify when the IoT-enabled device is malfunctioning, under attack from malware, or has the potential to destroy farms that they are supposed to properly irrigate for the effective and efficient growth of crops, it is crucial to classify the IoT-generated data from smart irrigation based on whether it is ON or OFF. Therefore, the inability of machine learning techniques to classify the data produced by these devices could indicate that they are unsuitable for use in smart irrigation. Four algorithms (LR, RF, SVM, and CNN) were selected for this study because they are the machine learning techniques that have been used the most frequently in the literature (Torres-Sanchez et al., 2020; Dhasaradhan et al., 2021). We evaluated our proposed approach using datasets from Mittal (2020).

2. RELATED WORKS

Automation technology is being used in agriculture more often to reduce costs and increase yields. Automation of irrigation systems is essential for controlling water and electricity usage in rice cultivation while preserving grain quality. To develop a rice irrigation system, Pfitscher et al. (2011) utilized wireless connectivity and Supervisory Control and Data Acquisition (SCADA). It demonstrated a completely automated irrigation system for rice. They used an ultrasonic sensor to monitor and control crop water levels, and a GPRS communication system to enable system remote control. Both water and power can be supplied by this gadget. By using soil moisture to determine when to turn ON or OFF the irrigation system, Tyagi et al. (2017) created an automated smart irrigation system. The developed system has benefits including minimal human intervention, great affordability, and simplicity of deployment (Tyagi et al., 2017).

On the basis of precise real-time farm data, Gondchawar and Kawitkar (2016) proposed a smart irrigation system with certain sensor modules for measuring light, soil moisture, humidity, and temperature. Furthermore, an automated watering system was developed by Shekhar et al. (2017) using Raspberry Pi 3 and Arduino Uno integrated electronics that are inexpensive. Shekhar et al. (2017) developed an intelligent IoT-based system that records and analyzes sensor data related to soil moisture and temperature for irrigation forecasting. This is an entirely automated system that uses irrigation intelligence and interacts with other devices. The Raspberry Pi 3 edge level processor receives the sensor data serially and employs the K-NN machine learning method to make an estimation of the soil quality based on a training set of data. In a direct connection to smart farming, several academics have argued for scalable network designs for managing and monitoring farms in rural areas (Monaco et al., 2016). Such systems enhance communication by utilizing technologies like the cross-layer-based channel access and routing solution for sensing and actuation to improve coverage range, throughput, and latency (Dholu & Ghodinde, 2018).

Reviewing the viability of using several proposed/developed automatic irrigation systems for rice irrigation were Kumar et al. (2018). The researchers chose rice because it requires a lot of water and is frequently produced underwater. According to the study, an autonomous irrigation system can make use of several cutting-edge irrigation approaches, including time-based systems, volume-based systems, real-time feedback mechanisms, and computer-based irrigation control systems (Kumar et al., 2018). A comprehensive IoT irrigation design was put up by Fernández-Ahumada et al. (2019) for the automation of irrigation networks for agricultural use. The developed system has shown to be extremely reliable and stable. A recent proposal for the detection and classification of intrusions on networks created and utilized for agricultural purposes was made by Raghuvanshi et al. (2022). For the purpose of classifying preprocessed data such as NSL, KDD, and NSL-KDD, they used machine learning techniques. These datasets are not generated by IoT smart irrigation systems, even though they are used for anomaly detection.

Iorliam et al. (2020) proposed a unique approach based on artificial neural networks for classifying soil data from IoT-enabled soil nutrients. Between 81.33% and 97.13% of accuracy was attained by their proposed approach. For the effective irrigation of crop areas, Roy et al. (2020) created an IoT automatic and manual irrigation system. Their research revealed that the AgriSens technology they developed increased crop output by 10.21% (Roy et al., 2020). Roy et al. (2020) proposed that in the future, machine learning should be used to analyze how factors like wind, humidity, temperature, and UV rays affect crop yields. Neforawati et al. (2019) classified the Paddy growth stage into seeding, transplanting, flowering, and maturity using CNN and found that their classifications had an overall accuracy of 82%. Also, Nindam et al. (2019) classified normal and abnormal Jasmine rice germination using CNN. According to experimental findings, their suggested method had a validation accuracy of 96.43%. Inspired by Iorliam et al. (2020), Roy et al. (2020), Neforawati et al. (2019) and Nindam et al. (2019), the researchers proposed using machine learning methods such as LR, RF, SVM, and CNN for the classification of IoT-enabled data generated while the smart irrigation devices are either ON or OFF.

3. PROPOSED METHOD

3.1. Dataset Description and Pre-Processing

As pointed out by Cheng et al. (2022), there is a general lack of public labelled IoT smart irrigation datasets for performing supervised learning experiments. This experiment therefore used publicly available data from Mittal (2020), which includes features like soil moisture (SM), temperature (T), humidity (H), and time (t). The algorithms employed in this experiment use each of these features as an input. The status column's "ON" or "OFF" values acted as class labels (ON = 1, OFF = 0).

The Kaggle dataset used to produce the smart irrigation IoT dataset is available at: http://autoirrigationdataforricecrop.herokuapp.com/. It is critical to classify the secondary data produced by IoT smart irrigation devices since it may include extremely intriguing statistical information that can characterize the device whether it is ON or OFF. Failure to classify these data points could indicate that the designed smart irrigation pump is failing or has been compromised by malware. Figure 1 displays the input data we used for this experiment, which includes columns for soil moisture (column A), temperature (column B), humidity (column C), time (column D), and status (column E).

	A	B	C	D	E
1	Soil Moisture	Temperature	Humidity	Time	Status
2	54	22	70	21	ON
3	12	20	40	104	OFF
4	34	26	35	62	ON
5	7	44	44	93	OFF
6	50	38	23	92	OFF
7	4	26	52	6	ON
8	15	34	58	82	ON
9	45	30	43	85	ON
10	47	4	42	2	OFF
11	19	41	22	0	ON
12	63	44	65	75	ON
13	88	34	41	71	ON
14	39	20	70	82	ON
15	65	43	44	93	OFF

Figure 1. *Input Data*

Machine learning algorithms (LR, RF, SVM, and CNN) were utilized to analyze and categorize the smart irrigation dataset based on the device status (ON = 1, OFF = 0). There are 100,000 samples in the overall dataset. 70% of the dataset is used for training, while 30% is used for testing. In a Google Colaboratory, Python is utilized for programming. The "dropna ()" method in Python is used to remove any rows with NULL values during data pre-processing.

3.2. Methodology

The proposed model gathers raw data from smart irrigation systems, pre-processes it to extract SM, T, H, and t, and then uses this processed data as input into four separate algorithms (LR, RF, SVM, CNN). The outcomes of these algorithms' classifications are subsequently produced as confusion matrices. Figure 2 summarizes the proposed model's functionality as a whole.

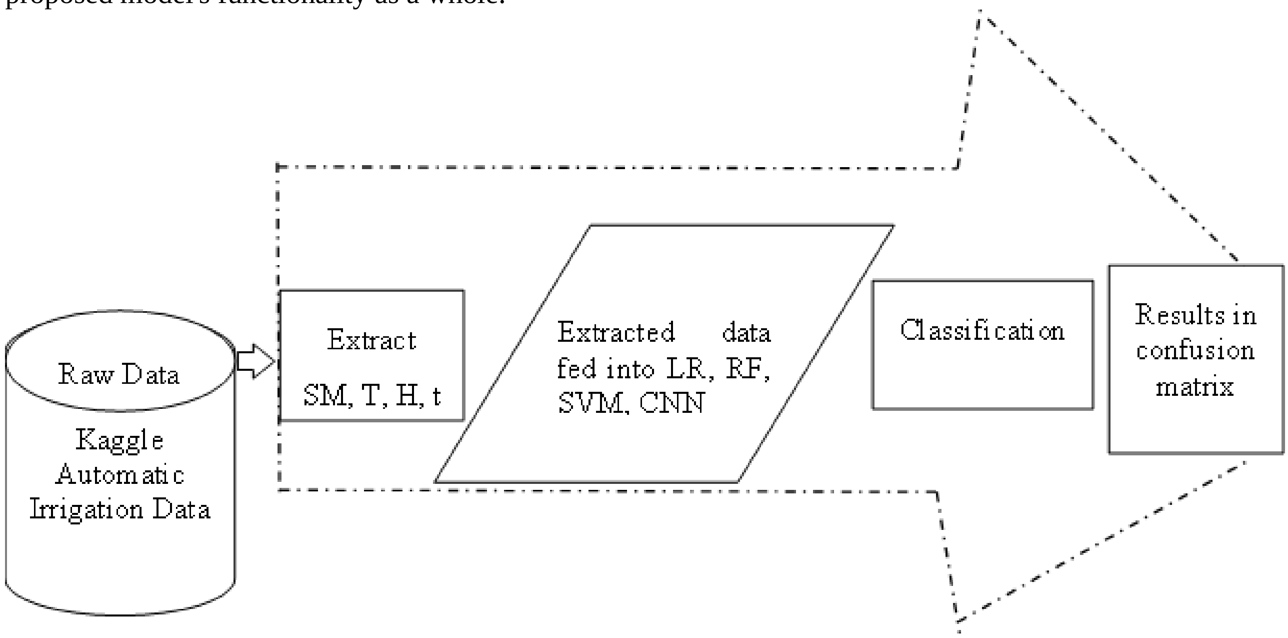

Figure 2. *Machine Learning Techniques for Classification of IOT-Enabled Smart Irrigation*

The following is a description of the machine learning techniques used in this paper:

i. Logistic regression (LR)

In this instance, the LR is used to classify data that is generated depending on whether the smart irrigation equipment is turned "ON" or "OFF." Since the LR technique uses supervised learning, labels are given for "ON = 1" and "OFF = 0". Either our dependent or target class is "ON" or it is "OFF." The variables that are independent are SM, T, H, and t.

Taking π as the probability of the event occurring, the LR model is therefore represented as:

$$\pi = P(Y = 1|X_1 \dots X_n = x_n) \tag{1}$$

$$= \frac{1}{1 + e^2} \tag{2}$$

where $z = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n$

and $\beta_0, \beta_1 \dots \beta_n$ are model parameters (estimators).

In our case, Y represents the outcome of our prediction (ON or OFF). $X_1, X_2, \dots, X_n$ are a set of n explanatory variables (SM, T, H, and t) (Bhowmik et al., 2011).

ii. Random forest (RF)

The RF has shown to improve the prediction performance of decision tree classifiers and controls over-fitting. As a result, it is chosen for the classification of data produced by smart irrigation devices.

The number of trees chosen for the forest (n-estimators) in our experiment was set to default = 100. The RF formula utilized for the classification task is based on the Gini index, which is mathematically stated as:

$$\text{Gini Index} = 1 - \sum_{k=1}^{n} (P_k)^2 \quad (3)$$

where P_k = the probability of an object being classified to a particular class (class 1 = “ON”, or class 0 = “OFF”) in the dataset

n= number of classes (2, ON or OFF) (Qi, 2012).

iii. Support vector machine (SVM)

Both separable and non-separable instances can be handled by the linear SVM. In our case, the radial basis function, or "RBF," was employed as the default kernel.

The following illustration shows a two-dimensional data set that can be divided along a line:

$$y = ax + b \quad (4)$$

if $y = x_2$, and $x = x_1$, the we get:

$$ax_1 - x_2 + b = 0 \quad (5)$$

When $x = (x_1, x_2), and\ w = (a, -1)$

We therefore get:

$$w * x_i + b = 0 \quad (6)$$

Based on the hyperplane, we can make predictions using the hypothesis function (h), defined as:

$$h(x_i) = \begin{cases} +1, & if\ w * x_i + b \geq 0 \\ -1, & if\ w * x_i + b < 0 \end{cases} \quad (7)$$

Therefore, any point above the hyperplane will be classified as +1 class, whereas any point below the hyperplane will be considered as -1 class (Fan, 2018).

iv. Convolutional Neural Network (CNN)

The CNN (Krizhevsky et al., 2017) is used in this paper since it has been demonstrated in literature to be one of the most often used artificial neural networks. The objective of this experiment is to use CNN to categorize IoT-smart irrigation data according to whether the device is "ON" or "OFF." The first class is denoted by the number 1 for "ON," and the second class, "class 0," is denoted by the number 0 for "OFF." The following steps illustrate the CNN setup used in this paper:

1. The dataset described in 3.1 is used
2. The sequential model "sequential ( )" is used as the first layer due to the fact that we are using linear stack of layers
3. Our model added the first hidden layer with 4 input parameters, and 480 neurons. The rectified linear activation function (ReLu) is first chosen due to its ability to achieve higher performance
4. Additional two dense layers are added with 240, and 120 neurons, respectively
5. Our model is ended with a dense layer, no activation, and a sigmoid activation function. The sigmoid activation function is chosen because we are considering a binary classification (ON or OFF) and the sigmoid activation function best fits our case in order to acquire our score
6. Our compiled model is done using the binary crossentropy, the Adam optimizer, and accuracy.
7. The training data of X, training data of Y are fit to our model using 1000 epochs, and a batch size of 128. The test data of X, and test data of Y are validated as well using the same epochs and batch size.

Based on the above setup, the results achieved are presented in Section 4 D.

3.3. Evaluation Metrics

To evaluate the proposed approach, the following evaluation measures are used:

i. **Accuracy**: Expressed mathematically as:

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{8}$$

ii. **Precision**: Expressed mathematically as:

$$\text{Precision} = \frac{(\text{TP})}{(\text{TP} + \text{FP})} \tag{9}$$

iii. **Recall**: Expressed mathematically as:

$$\text{Recall} = \frac{(\text{TP})}{(\text{TP} + \text{FN})} \tag{10}$$

iv. **F1-Score**: Expressed mathematically expressed as:

$$\text{F1} - \text{Score} = \frac{(2 \text{ x Precision x Recall})}{(\text{Precision} + \text{Recall})} \tag{11}$$

where:

True Positive (TP): The outcome of the developed model correctly predicts the positive class.

True Negative (TN): The outcome where the negative class is correctly predicted by the developed model.

False Positive (FP): The outcome where the positive class is incorrectly predicted by the developed model.

False Negative (FN): The outcome where the negative class is incorrectly predicted by the developed model.

4. RESULTS AND DISCUSSIONS

To classify data from IoT-enabled smart irrigation devices, four algorithms were used (LR, RF, SVM, and CNN). Based on the information produced by the IoT smart irrigation system, the machine learning algorithms were able to divide the dataset into two classes. Class 0 (information while the smart irrigation pump is off) and Class 1 (data when the smart irrigation pump is ON). The data (100,000 rows) are divided at random into 70% training and 30% testing, with the results displayed and discussed in 4 A–D.

A. Performance of Logistic Regression

The pre-processed IoT-enabled smart irrigation data given in 3.1 are supplied into the LR algorithm. As can be seen in Figure 3, the logistic regression technique achieved 71.76% accuracy, 0.7155 F1 score, 0.7303 precision value, and 0.7501 recall, respectively.

```
***LOGISTIC REGRESSION RESULTS***
Accuracy: 71.76%
F1 Score: 0.715479
Precision:  0.7302731181493369
Recall:  0.7500777508241587
```

Figure 3. *Results of Logistic Regression*

Again, Figure 4 shows the result of the confusion matrix for the LR algorithm. The LR algorithm correctly classified 9469 as class 0 (OFF) and wrongly classified 4454 samples of class 0 as class 1 (ON). Furthermore, the LR algorithm correctly classified 12059 samples of class 1 (ON), and wrongly classified 4018 samples of class 1 as class 0 (OFF).

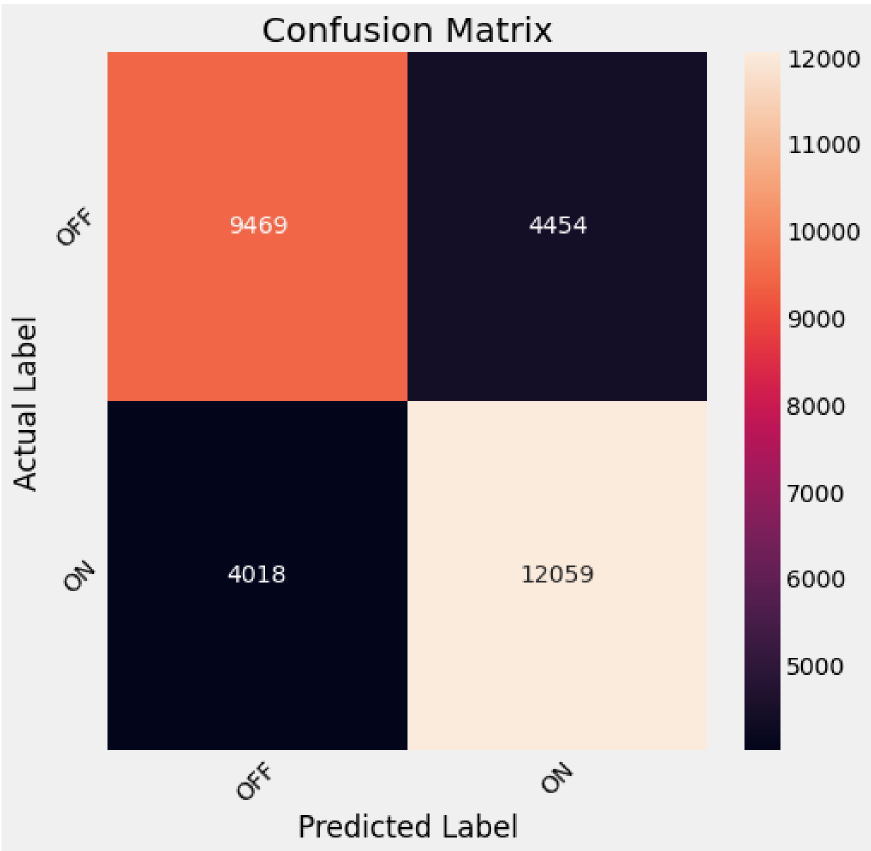

Figure 4. *Confusion Matrix for Logistic Regression*

Recent research by Raghuvanshi et al. (2022) showed that the LR's classification accuracy on intrusion detection datasets like NSL, KDD, and NSL-KDD was less than 78%. Our classification accuracy using LR on a realistic smart irrigation dataset is 71.76%, which is quite similar to the outcome in Raghuvanshi et al. (2022).

B. Performance of Random Forest

As can be observed in Figure 5, the Random Forest (RF) algorithm achieved 99.98%, 0.9998, 0.9996, and 0.9999 for the Accuracy, F1 score, Precision value, and Recall, respectively.

```
***RANDOM FOREST RESULTS***
Accuracy: 99.98%
F1 Score: 0.999765
Precision:  0.9996269120756125
Recall:  0.9999377993406731
```

Figure 5. *Results for Random Forest*

The RF's confusion matrix shows that the algorithm correctly classified 13917 as class 0 (OFF) and wrongly classified 6 samples of class 0 as class 1 (ON). Furthermore, the RF algorithm correctly classified 16076 samples of class 1 (ON), and wrongly classified only 1 sample of class 1 as class 0 (OFF) as seen in Figure 6.

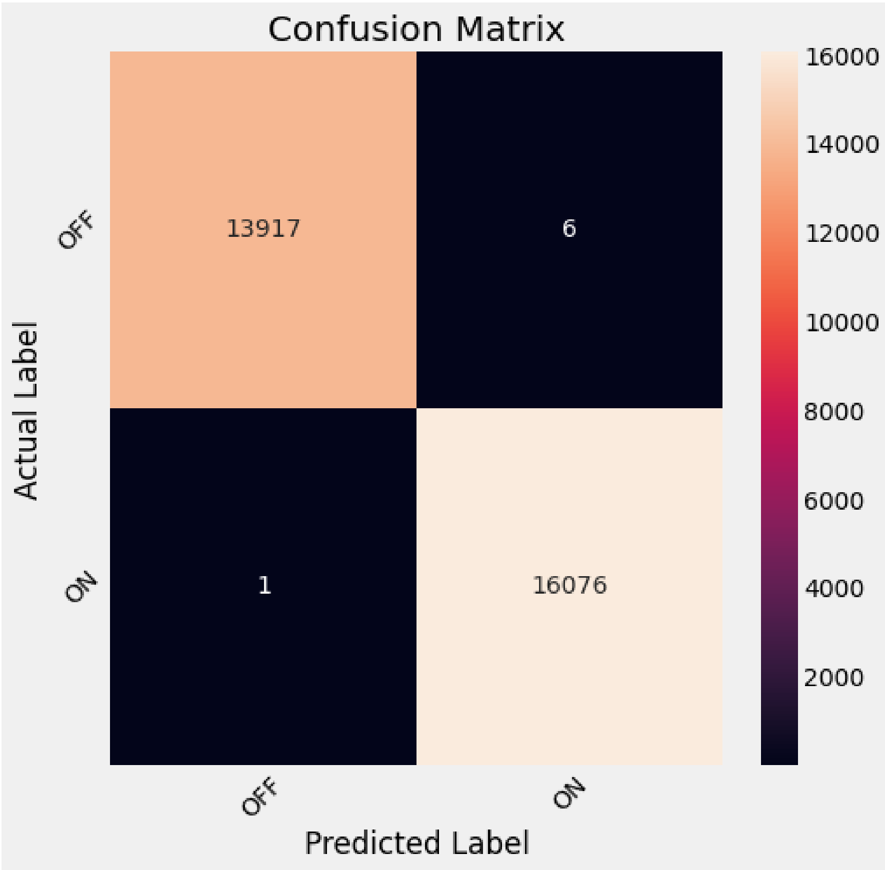

Figure 6. *Confusion Matrix for Random Forest*

This shows that the RF algorithm accurately classified data generated by IOT smart irrigation device up to 99.98%. Our accuracy outperformed that of Raghuvanshi et al. (2022), when they used the RF algorithm on intrusion detection datasets such as NSL, KDD, and NSL-KDD and achieved an accuracy of less than 78% (Raghuvanshi et al., 2022). Moreover, our accuracy also outperformed that of Ok et al. (2012) when they utilized the random forest algorithm for agricultural crop classification and achieved an accuracy of 85.89%.

C. Performance of Support Vector Machine

As shown in Figure 7, the Support Vector Machine (SVM) algorithm achieved 90.21% Accuracy, 0.9014 F1 scores, 0.9019 Precision value, and 0.9170 Recall.

```
***SVM RESULTS***
Accuracy: 90.21%
F1 Score: 0.901423
Precision:  0.9018780204318835
Recall:  0.9170243204577968
```

Figure 7. *Support Vector Machine Results*

Again, the confusion matrix for the SVM shows that it correctly classified 12319 samples as class 0 (OFF) and wrongly classified 1604 samples of class 0 as class 1 (ON). Furthermore, the SVM algorithm correctly classified 14743 samples of class 1 (ON), and wrongly classified 1334 samples of class 1 as class 0 (OFF) as seen in Figure 8.

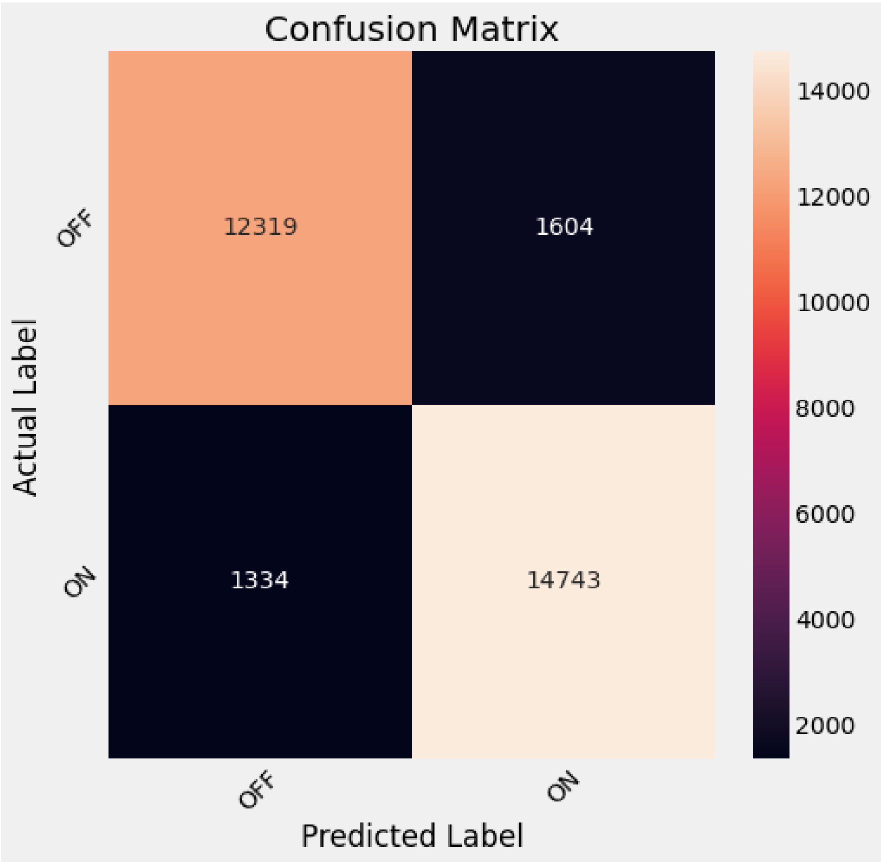

Figure 8. *Confusion Matrix for Support Vector Machine*

This further proves that our proposed method can effectively be applicable for the separability of smart irrigation IoT data using SVM with a high accuracy of 90.21%. Our proposed method performs a bit lesser as compared to Raghuvanshi et al. (2022), when they used the SVM algorithm on intrusion detection datasets such as NSL, KDD, and NSL-KDD and achieved an accuracy of more than 98% (Raghuvanshi et al., 2022). Interestingly, our proposed method achieved a slightly higher result to that of Jagtap et al. (2022) when they applied the SVM in the classification of crop images and achieved an accuracy that was less than 90%.

D. Performance of Convolutional Neural Network (CNN)

The Convolutional Neural Network (CNN) algorithm achieved an accuracy of 0.9823, precision of 0.98, recall of 0.98, and F1 score of 0.98. Furthermore, the training loss vs the epochs for the CNN classification of the IoT-enabled dataset is presented in Figure 9.

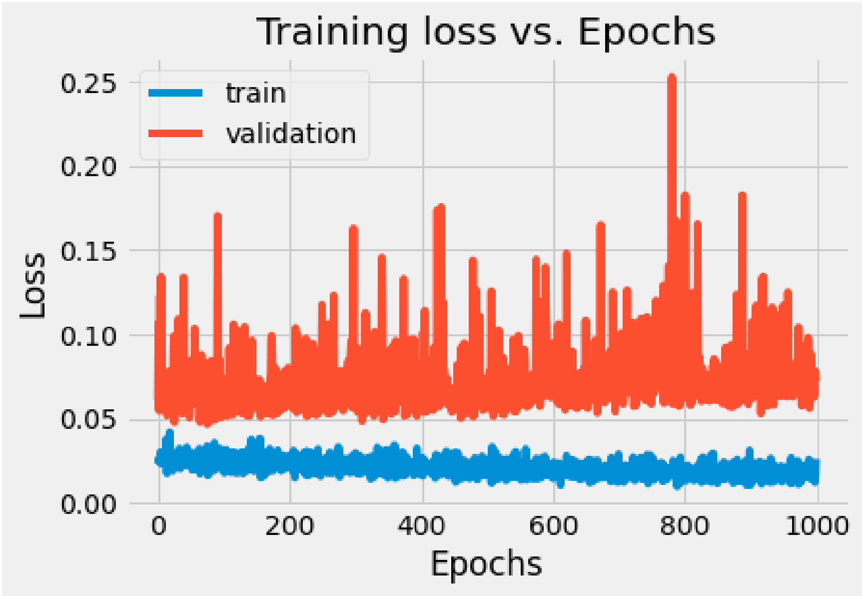

Figure 9. *Training Loss Vs Epochs for the CNN*

Generally, when the Loss Vs Epochs graph is taken into account, lower Loss values result in a good performance. Figure 9 shows that the graph was more consistent for the train datasets. The Loss values got closer to zero as the number of Epochs tends towards 1000. Additionally, as Epoch values tends towards 1000, the Loss values tends more towards zero (0) for the validation dataset.

Once more, while examining the training accuracy vs. epochs for this experiment as shown in Figure 10, it can be seen that the accuracies for the train dataset and test dataset were over 0.98 when the epochs were closer to 1000.

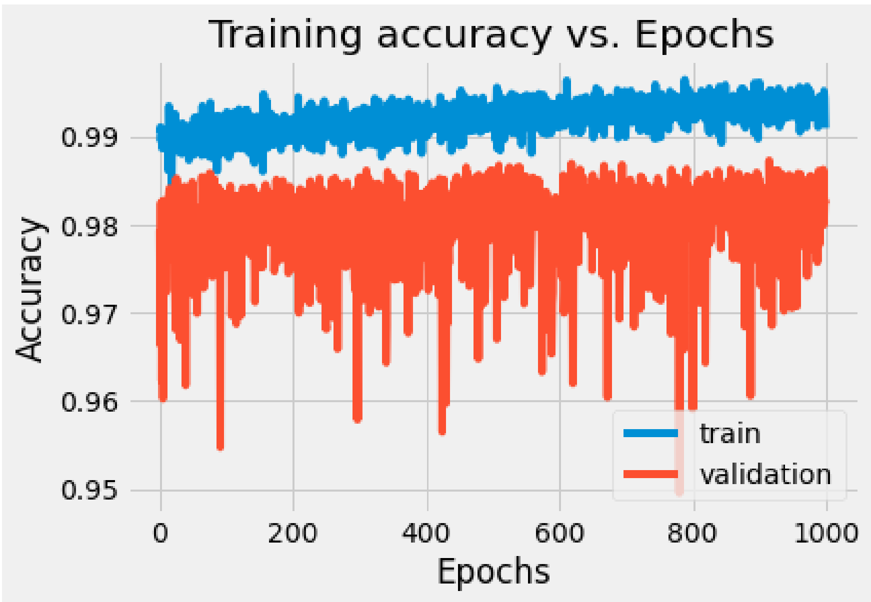

Figure 10. *Training Accuracy Vs Epochs for the CNN*

This demonstrates that for this experiment, epochs of over 1000 were well suited for the effective classification of IoT enabled data.

The confusion matrix is shown in Figure 11 to clearly demonstrate the number of samples that were classified using CNN based on the "ON" or "OFF" of the smart irrigation device.

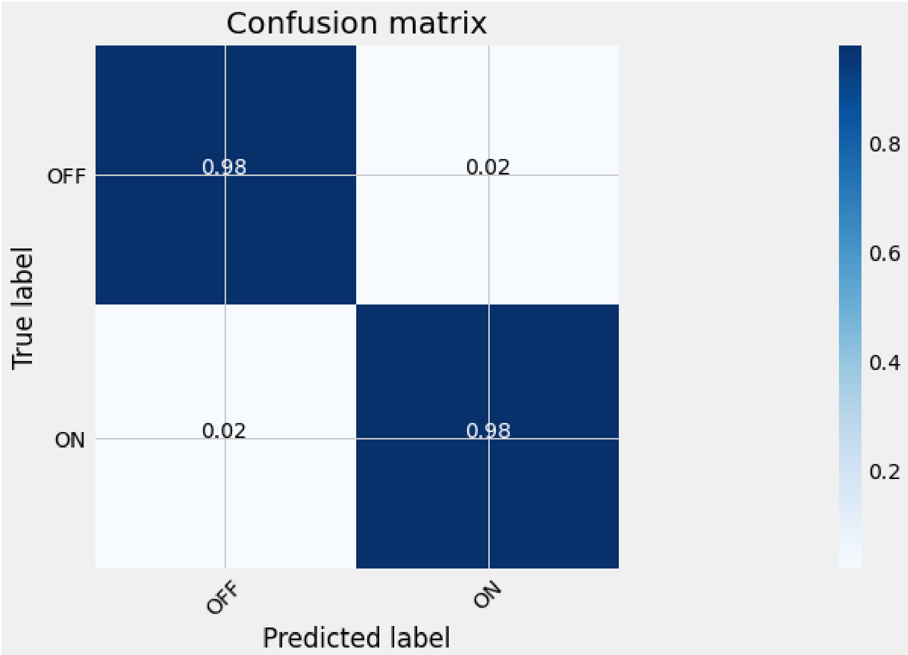

Figure 11. *Confusion Matrix for the CNN*

It can be seen from Figure 11 that the CNN algorithm correctly classified 0.98 (98%) samples as class 0 (OFF) and wrongly classified 0.02 (2%) samples of class 0 as class 1 (ON). Furthermore, the CNN algorithm correctly classified 0.98 (98%) samples of class 1 (ON), and wrongly classified 0.02 (2%) samples of class 1 as class 0 (OFF). With an accuracy of up to 98% shown in the confusion matrix, it proved that the proposed method of using the CNN for the classification of IoT-enabled data was very successful.

The accuracy of our proposed method (98.23%) outperformed that of Neforawati et al. (2019), when they used the CNN to classify Paddy growth level and achieved an overall accuracy of 82%. Furthermore, the accuracy of our proposed method outperformed that of Nindam et al. (2019) when they utilized the CNN for the classification of normal and abnormal Jasmine rice germination and achieved a validation accuracy of 96.43%.

In summary, Table 1 and Figure 12 compares the results of the machine learning techniques achieved in this paper.

Table 1. *Comparison Performance of Algorithms*

Algorithm	**Accuracy**	**F1 score**	**Precision Value**	**Recall**
LR	71.760	0.715	0.730	0.750
RF	99.980	0.999	0.999	0.999
SVM	90.210	0.901	0.901	0.917
CNN	98.230	0.980	0.980	0.980

Applying machine learning techniques for the separability of data produced by IoT-enabled smart irrigation devices is necessary because, if an IoT smart irrigation pump is supposed to go "OFF" but does not, this could result in flooding of the farm that is receiving irrigation, and vice versa. Therefore, the precise classification of data derived from whether the smart irrigation pump is "OFF" or "ON" can help in identifying when such a device is faulty (malfunctioning), under attack from malware, and has a potential of destroying farms which they are supposed to properly irrigate for the effective and efficient growth of crops.

In this study, the RF algorithm outperformed the SVM, LR, and CNN in all the evaluation metrics employed. The classification accuracy of the RF is 99.98%, followed by the CNN which is 98.23%, then the SVM which is 90.21%, and the LR which is 71.76%. From the comparative analysis, it can also be deduced that the LR performed least in all evaluation metrics used in this study.

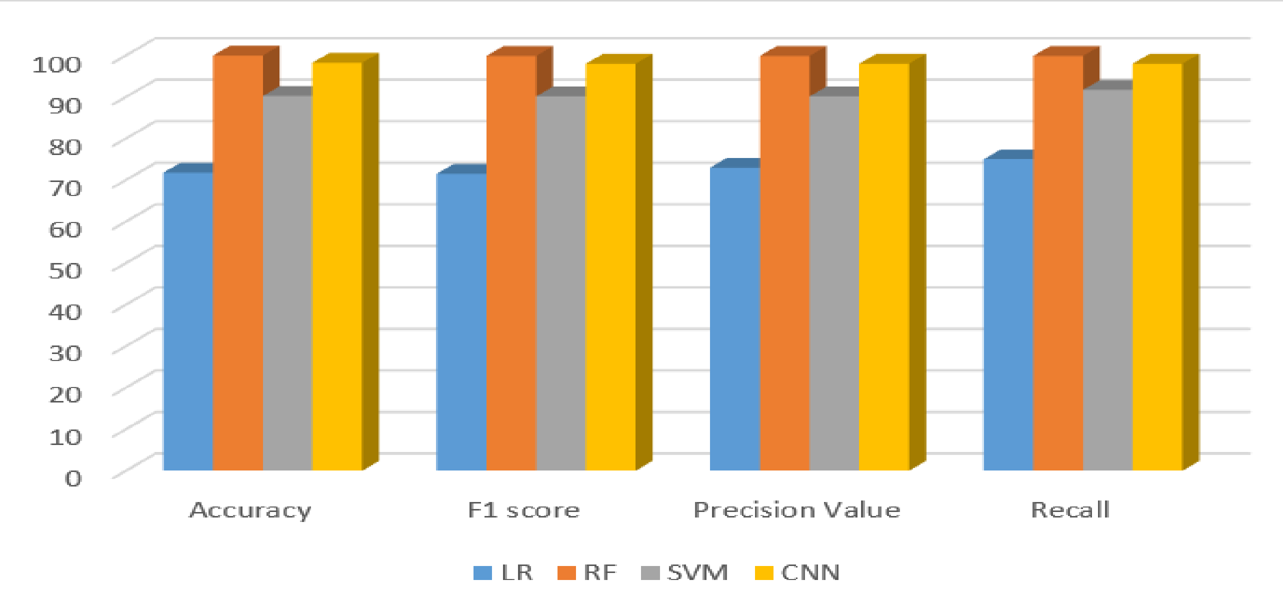

Figure 12. *Graph Showing the Performance of Algorithms*

5. CONCLUSION AND FUTURE WORK

This study examines how to classify (separate) data produced by IoT-enabled smart irrigation devices using machine learning approaches. The researchers concluded that smart irrigation device pre-processed IoT-generated datasets can be employed as input data into machine learning algorithms to understand patterns and carry out classification tasks. The RF algorithm, which was one of the four machine learning algorithms (LR, RF, SVM, and CNN) used in this study, was judged to be the most successful because it accurately distinguished between the times when the smart irrigation device was "ON" and "OFF" with 99.98%, 0.9998, 0.9996, and 0.9999 for the Accuracy, F1 score, Precision value, and Recall, respectively. The precise classification of IoT smart irrigation data by machine learning algorithms may indicate that the IoT devices that produced such data are particularly effective and efficient in the irrigation of crop farms, which may result in improved crop yields. In the future, we intend to classify IoT-enabled smart irrigation data for agricultural purposes using Benford's law and Zipf's law.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Adam, M. S. A., Osman, A. A., Omer, E. A., & Abdallah, A. M. B. (2020). *Automatic Irrigation Implementation*. PhD Thesis, Supervised by Ust. Jafer Babiker, Sudan University of Science & Technology.

Bhowmik, A., Ramasubramanian, V., & Kumar, A. (2011). Logistic regression for classification in agricultural ergonomics. *Advances in Applied Science Research*, *3*(2):163-170.

Çetin, M., & Beyhan, S. (2022). Smart Irrigation Systems Using Machine Learning and Control Theory. In: R. Bhatnagar, N. K. Tripathi, N. Bhatnagar, & C. K. Panda (Eds.), *The Digital Agricultural Revolution: Innovations and Challenges in Agriculture through Technology Disruptions* (pp. 57-85). Scrivener Publishing LLC. doi:10.1002/9781119823469.ch3

Cheng, W., Ma, T., Wang, X., & Wang, G. (2022). Anomaly Detection for Internet of Things Time Series Data Using Generative Adversarial Networks with Attention Mechanism in Smart Agriculture. *Frontiers in Plant Science*, *13*. doi:10.3389/fpls.2022.890563

Dhasaradhan, K., Jaichandran, R., Shunmuganathan, K. L., Usha Kiruthika, S., & Rajaprakash, S. (2021). Hybrid machine learning model using decision tree and support vector machine for diabetes identification. In: V. Bhateja, S. C. Satapathy, C. M. Travieso-González, V. N. M. Aradhya (Eds.), *Data Engineering and Intelligent Computing* (Proceedings of ICICC 2020) (pp. 293-305). Springer. doi:10.1007/978-981-16-0171-2_28

Dholu, M., & Ghodinde, K. A. (2018, May). *Internet of things (IoT) for precision agriculture application*. In: Proceedings of the 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 339-342). IEEE. doi:10.1109/ICOEI.2018.8553720

Fan, S. (2018, May 7). Understanding the mathematics behind Support Vector Machines. (Accessed: 30/06/2022) URL

Fernández-Ahumada, L. M., Ramírez-Faz, J., Torres-Romero, M., & López-Luque, R. (2019). Proposal for the design of monitoring and operating irrigation networks based on IoT, cloud computing and free hardware technologies. *Sensors*, *19*(10), 2318. doi:10.3390/s19102318

Goap, A., Sharma, D., Shukla, A. K., & Krishna, C. R. (2018). An IoT based smart irrigation management system using Machine learning and open source technologies. *Computers and Electronics in Agriculture*, *155*, 41-49. doi:10.1016/j.compag.2018.09.040

Gondchawar, N., & Kawitkar, R. S. (2016). IoT based smart agriculture. *International Journal of Advanced Research in Computer and Communication Engineering*, *5*(6), 838-842.

Iorliam, A., Adeyelu, A., Otor, S., Okpe, I., & Iorliam, I. B. (2020). A Novel Classification of IoT-Enabled Soil Nutrients Data Using Artificial Neural Networks. *International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering*, 8(4), 103-109. doi:10.17148/IJIREEICE.2020.8418

Iorliam, A., Iorliam, I. B., & Blum, S. (2021). Internet of Things for Smart Agriculture in Nigeria and Africa: A Review. *International Journal of Latest Technology in Engineering, Management & Applied Science*, *10*(2), 7-13.

Jagtap, S. T., Phasinam, K., Kassanuk, T., Jha, S. S., Ghosh, T., & Thakar, C. M. (2022). Towards application of various machine learning techniques in agriculture. *Materials Today: Proceedings*, *51*(1), 793-797. doi:10.1016/j.matpr.2021.06.236

Janani, M., & Jebakumar, R. (2019). A study on smart irrigation using machine learning. *Cell & Cellular Life Sciences Journal*, *4*(1), 1-8. doi:10.23880/cclsj-16000141

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84-90. doi:10.1145/3065386

Kumar, M., Sahni, R. K., Waghaye, A. M., Nayak, A. K., & Kumar, D. (2018). Automated Irrigation System for Rice: A Review. *The Andhra Agric. Journal*, *65 (spl)*, 324-329.

Mittal, R. (2020). Automatic Irrigation Data for Rice Crop. (Accessed: 30/06/2022) URL

Monaco, F., Sali, G., Ben Hassen, M., Facchi, A., Romani, M., & Valè, G. (2016). Water management options for rice cultivation in a temperate area: A multi-objective model to explore economic and water saving results. *Water*, *8*(8), 336. doi:10.3390/w8080336

Nawandar, N. K., & Satpute, V. R. (2019). IoT based low cost and intelligent module for smart irrigation system. *Computers and Electronics in Agriculture*, *162*, 979-990. doi:10.1016/j.compag.2019.05.027

Neforawati, I., Herman, N. S., & Mohd, O. (2019, April). Precision agriculture classification using convolutional neural networks for paddy growth level. *Journal of Physics: Conference Series*, *1193*, 012026. doi:10.1088/1742-6596/1193/1/012026

Nindam, S., Sung, T. L., Manmai, T.-O., & Lee, H. J. (2019, June). *Collection and Classification of Jasmine Rice Germination Using Convolutional Neural Networks*. In: Proc. International Symposium on Information Technology Convergence (ISITC 2019) (pp. 105-108).

Ok, A. O., Akar, O., & Gungor, O. (2012). Evaluation of random forest method for agricultural crop classification. *European Journal of Remote Sensing*, *45*(1), 421-432. doi:10.5721/EuJRS20124535

Pfitscher, L. L., Bernardon, D. P., Kopp, L. M., Ferreira, A. A. B., Heckler, M. V. T., Thome, B. A., Montani, P. D. B., & Fagundes, D. R. (2011, May). *An automated irrigation system for rice cropping with remote supervision*. In: J. A. Aguado, & A. Pires (Eds.), 2011 International Conference on Power Engineering, Energy and Electrical Drives. 2011 International Conference on Power Engineering, Energy and Electrical Drives (POWERENG). IEEE. doi:10.1109/PowerEng.2011.6036452

Pluchinotta, I., Pagano, A., Giordano, R., & Tsoukiàs, A. (2018). A system dynamics model for supporting decision-makers in irrigation water management. *Journal of Environmental Management*, *223*, 815-824. doi:10.1016/j.jenvman.2018.06.083

Qi, Y. (2012). Random Forest for Bioinformatics. In: C. Zhang, & Y. Ma (Eds.), *Ensemble Machine Learning Methods and Application* (pp. 307-323). Springer, Boston, MA. doi:10.1007/978-1-4419-9326-7_11

Raghuvanshi, A., Singh, U. K., Sajja, G. S., Pallathadka, H., Asenso, E., Kamal, M., Singh, A., & Phasinam, K. (2022). Intrusion Detection Using Machine Learning for Risk Mitigation in IoT-Enabled Smart Irrigation in Smart Farming. In: M. F. Manzoor, A. Hussain, & R. M. Aadil (Eds.), *Journal of Food Quality*, *2022 (Special Issue)*, 3955514. doi:10.1155/2022/3955514

Romero, M., Luo, Y., Su, B., & Fuentes, S. (2018). Vineyard water status estimation using multispectral imagery from an UAV platform and machine learning algorithms for irrigation scheduling management. *Computers and Electronics in Agriculture*, *147*, 109-117. doi:10.1016/j.compag.2018.02.013

Roy, S. K., Misra, S., Raghuwanshi, N. S., & Das, S. K. (2020). AgriSens: IoT-based dynamic irrigation scheduling system for water management of irrigated crops. *IEEE Internet of Things Journal*, *8*(6), 5023-5030. doi:10.1109/JIOT.2020.3036126

Sanjeevi, P., Prasanna, S., Siva Kumar, B., Gunasekaran, G., Alagiri, I., & Vijay Anand, R. (2020). Precision agriculture and farming using Internet of Things based on wireless sensor network. *Transactions on Emerging Telecommunications Technologies*, *31*(12), e3978. doi:10.1002/ett.3978

Shekhar, Y., Dagur, E., Mishra, S., Tom, R. J., Veeramanikandan, M., & Sankaranarayanan, S. (2017). Intelligent IoT based automated irrigation system. *International Journal of Applied Engineering Research*, *12*(18), 7306-7320.

Surendran, U., Sushanth, C. M., Mammen, G., & Joseph, E. J. (2015). Modelling the crop water requirement using FAO-CROPWAT and assessment of water resources for sustainable water resource management: A case study in Palakkad district of humid tropical Kerala, India. *Aquatic Procedia*, *4*, 1211-1219. doi:10.1016/j.aqpro.2015.02.154

Torres-Sanchez, R., Navarro-Hellin, H., Guillamon-Frutos, A., San-Segundo, R., Ruiz-Abellón, M. C., & Domingo-Miguel, R. (2020). A decision support system for irrigation management: Analysis and implementation of different learning techniques. *Water*, *12*(2), 548. doi:10.3390/w12020548

Tyagi, A., Gupta, N., Navani, J. P., Tiwari, R., & Gupta, A. (2017). Smart irrigation system. *International Journal for Innovative Research in Science & Technology*, *3*(10).