International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

IoT Enabled WSN and Machine Learning Techniques to Surveillance the Smart Irrigation System

Akshatha Y[1], Dr. A S Poornima*[2]

Submitted: 25/04/2023 **Revised**: 27/06/2023 **Accepted**: 07/07/2023

Abstract*:* Farmers try to follow their experience in farming. They will not be aware of the changes that happens in the environment that directly affect the profit as well as the outcome from the farm field. The adoption of smart technologies in farming helps the farmers to utilize resources efficiently. The important resource of farming is water. The consumption of water should be frugal enough to regulate the water supply. The incompetent way of irrigation results in water wastage. Therefore, implementing IoT-enabled WSN techniques allows to moderate the irrigation system known as smart farming or smart agriculture. The smart irrigation system helps to efficiently use the water resource in farming. Sensors are deployed to sense environmental conditions such as temperature, humidity and soil moisture content. The continuous sensed data will be processed and sent to farmers. Based on the data received, farmers decide the water supply to the farm field. Injecting false or incorrect data results in overwatering or under-watering that affects the yield of the farm field directly. Detection of such false data is accomplished using two phases of classification. The two phases of classifications implemented in the proposed scheme are machine learning techniques and fault detection algorithm (FDA). The data that are classified as non-anomalous from the first phase of classification are subjected to the second phase that is FDA. This two-phase of classification concludes that the data received are non-anomalous. This entire classification process is done at the base station (BS). The data that are detected as anomalous either at the first phase or at the second phase are dropped directly without considering for further process. The data that are processed completely at BS will be forwarded to the farmers of the particular farm field.

Keywords: *IoT, Machine Learning Techniques, Smart Irrigation, Smart Farming*

1. Introduction

The foundation and important industry of Indian economy is agriculture. Automation of agriculture is the major concern and is one of the emerging concept in almost all the countries. The population in all the countries are increasing expeditiously, as there is a growth in the world's population, there increases the demand for food. The demand for food as well as the customer's demand have made famers not to allow that satisfy the requirement [1][2][3]. These requirements made difficult for the famers to improvise the traditional techniques and practices. Agriculture us one of the most important sector in present society. Hence it is essential to guarantee that enhancement are done in the agriculture sector to upgrade its production and the outcomes. It is difficult to balance the agricultural resources because almost all the courtiers depends majorly on agriculture sector [4]. Water is one of the major resource in agriculture field. Adopting smart technologies into agriculture improves the traditional practice into modern generation of agricultural practice. Smart technology or smart computing is majorly dependent on IoT [5][6][7][8][9]. IoT is a network composed of physical objects or things that can communicate with one another to exchange the data. Researchers in entire world are trying to implement IoT technologies everywhere to make everything smart such as smart cities, smart healthcare, smart farming [6][7][10][11][12][13]. Especially, in the field of agriculture, adopting IoT enabled WSN [14] kind of smart technologies helps the farmers to get improvement which relay on the country's economic growth as well as fulfils the demand of food in the world. This also helps in the situation where famers cannot be available in all 24 hours in the farm field and he will not be knowing of operating the devices to calculate the environment condition. The measurements of such environmental conditions help the farmers to monitor the field without his intervention. The automatic system practice in the agriculture field runs without human intervention and the system will notify the famers regarding the changes that happens in the farm field to make the decision if any critical problems occurs that are to be solved immediately. Such kind of farming practice allows the farmers to maintain the farmer fields that widely spread across the large area.

1 Department of Computer Science and Engineering, SIT, Tumakuru, Karnataka, INDIA
ORCID ID : 0000-0002-1585-2308
2 Department of Computer Science and Engineering, SIT, Tumakuru, Karnataka, INDIA
ORCID ID : 0000-0001-9458-4251
** Corresponding Author Email: akshathay@sit.ac.in*

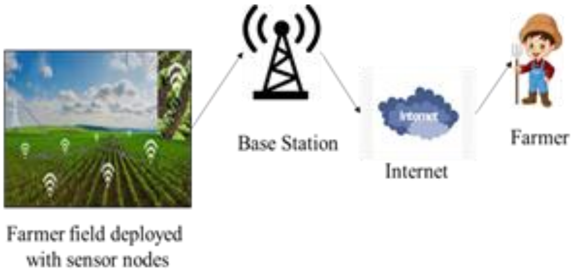

Fig. 1: The overall model of the Farm field.

Figure 1 depicts the overall model of the farm field which uses the smart technologies. The figure explains the sensor nodes are deployed in the farm field to measure environmental condition which is generally wireless sensor network. Respective sensor nodes are used to measure the respective parameters. The sensed data are sent to base station (BS) and through the internet sensed data will be forwarded to farmers. The sensed data are influential as it used to improvise the farming. The scenarios we make use of the sensed data are as follows: controlling irrigation by sensing moisture level in the soil, releasing the manure periodically to get better yield of crops etc, when making such important decision by using sensed data, its misuse may result in disaster in the farm field.

In this paper, we propose secure and smart irrigation system using machine learning and WSN enabled IoT system. The application we spotted light here is the large farming area, which is deployed with sensor nodes randomly all over the farm to sense temperature, humidity and soil moisture content. The collected data used to decide whether water supply is required or not. The data sensed are sent to BS. The job of BS is the aggregate the collected data.

Based on the aggregated data, BS decides whether to release the water or stop the water. In the smart irrigation scenario, if an attacker injects the false data into the considered parameters, it may result in overwatering or under-watering which effects on the crops in the farm field. We propose her two level of classification method to classify the receive data from the sensor as anomalous or non-anomalous. Only the data that belongs to non-anomalous are used in aggregation process which involves in making the decision of irrigation system. The two phases of data classification are as shown, in the first phase machine learning models are used to classify the data as anomalous or non-anomalous. The data classified as non-anomalous will involve second phase of classification by using fault detection algorithm (FDA). For experimentation purpose, we considered 3500 records. From the single dataset three different datasets were generated by varying anomalous data in each dataset. The dataset that are considered for classification are as follows,

In the first dataset, 10% of anomalous data is added. In the second dataset, 20% of anomalous data is added. In the third dataset, 10% extra added to second dataset. During experimentation for all the datasets, the dataset is split into 70:30 ratio ie, we considered 70% of dataset for training the model and remaining 30% of data is used to test the model. Machine learning technologies used are K-Nearest Neighbor (K-NN), Decision tree, Logistic Regression, Naïve Bayes, Support vector Machine (SVM) and multi-layer perceptron. Among all the three datasets with varying anomalous records, decision tree algorithm gives an accuracy of 91%, K-NN, SVM, logistic regression gives the better accuracy of 86% and Naïve Bayes gives 73% of accuracy compared to other algorithms. Therefore, the decision tree is used as the first phase of classifier. Only the data that are classified as non-anomalous in the first phase are subjected to FDA which is second phase of classification.

Smart farming is used to monitor the farm field with the help of respective sensors, this technique helps the farmers to look after the farm field from anywhere.

Smart agriculture using IoT

The automated irrigation system is presented in [5]. The authors use Node microcontroller (NodeMCU) to achieve the objectives. NodeMCU is an open-source hardware and software development environment. The method supplies the water to the plants based on the soil moisture sensor reading.

The implementation of automatic soil moisture monitoring system with the integration of WSN and Bluetooth is presented in [7]. The Bluetooth is used to transmit and receive data from sender to receiver, monitor the soil, save data, the variation of water is displayed, and neural network is used to improve the performance of method used for controlling water. Different types of sensors are used for sensing respective data. Based on these sensor values releasing of fertilizers and other agriculture chemicals required for the farm field.

Because of issues raising in the recent years, the scheme [3] gives the review on the application of WSN in the irrigation field. The paper concludes that, efficient and optimum usage of water assurances and contributes the reduction in minimizing security of water crises.

To help the farmers who the farmers who wish to contribute to country, [5] gives solution. Since availability of farmers in the farm field is not possible at every time, hence to allow surveillance on personnel and to avoid crops from losses, the authors contributed the method which is easy to use from smart phones.

Security requirement in IoT networks

Today, whole world is depended on new emerging technologies. Hence we are surrounded with number of smart devices. These smart devices making our routine easy and convenient the smart devices are exposed to many threats and cyber-attacks [6][15][16][17][18][19]. The major security concern in IoT network are, integrity, confidentiality, authentication and data availability. The major threats of WSN are, physical attack, node replication, selective forwarding, worm hole attack, Sybil attack, sink hole attack, eavesdropping. According to the authors [6], to obtain secure communication between devices in IoT, it is required to involve encryption, end-to-end environments, access control and infrastructure protection.

Smart Agriculture with security

Since farming is more influenced by new technologies, it is vulnerable to more attacks. Hence it is necessary to protect from the attacker.

To deal with the current security issues, [8] gives the survey on attacks that destroy the security aspects in the field of smart farming. Cyber threats also highlighted in this paper. The involvement of new technologies inherits several security challenges that occur raises the underlying consciousness regarding empirical methodology. Smart farming shows interest on the IoT application [10][11][12][20][21][22] that monitors the environment remotely. The information that communicates in farm field are unprotected. In the smart farming technique, the model automatically keeps and manages the information about temperature, humidity. An attacker can gain the vulnerabilities of security [12][23][24] to maintain the temperature, humidity which may cause serious issue. In [10] an authentication method has been proposed to perform with minimum encryption and decryption steps. In this scheme, the session key is combined with public key to control the farm securely. The advantage of the scheme is to reduce the encryption/decryption time, time used for registration also reduced, and mainly the scheme eases the use of the smart cards which is having less computational performance. The computation performance is less in the scheme by the use of session key when compared to available authentication problems in smart field.

The above stated models are homogenous with respect to the smart farming using IoT [25][26][27]. The models helps to automate the irrigation system, water control automatically, the scheme summarizes the security challenges faced in the farm field. The proposed scheme helps to classify the anomalous or non- anomalous data which is received from sensor node using two phase verification process.

2. Materials and Methods

Here we propose, secure and smart irrigation system. The main aim of the proposed scheme is to detect the anomalous or non-anomalous data in the considered dataset. The automatic irrigation system works on the basis of data collected from various sensor that are deployed in the farm field. Each sensor node S_i senses the parameter at each interval of time t and send the same to BS. The BS aggregates the data that are received to decide whether the water supply is necessary or not. Since water is the important resource in agriculture, wastage of water resource effects on crop production. There occurs a way to lose the water resource that is, if an intruder injects the false data of sensed parameters. It results in under-watering or overwatering. Hence it is necessary to check the anomalous data in the received dataset before start computing the aggregate. There is a necessity of system that is able to distinguish the anomalous data from the dataset received. Here two phases are proposed to classify the data

Phase 1: Using machine learning techniques or models to classify the received as either anomalous or non-anomalous.

Phase 2: The data that are classified as non-anomalous there invokes second phase of security. For second phase of data classification, FDA is used.

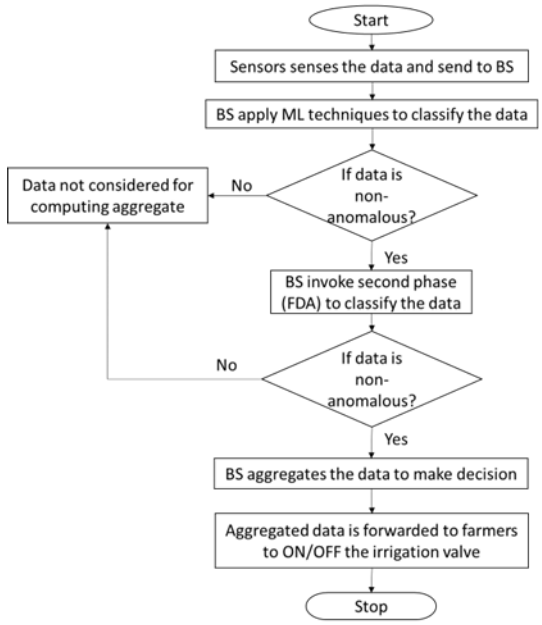

Fig. 2. The Flow diagram of the proposed scheme

The flow diagram of the proposed scheme are shown in figure 2.

Figure 2:

The data sensed from sensor node gives the data regarding the real time environmental factor like temperature, humidity, and soil moisture. These parameters are used to operate the irrigation system. Before the starting the aggregation process, the BS first classify the data as anomalous or non-anomalous. The BS itself impossible to classify the data. Hence it is important to consider the data probability which leads to better transfer of the data within the network. The anomalous data transfer leads to either overwatering or under-watering. Hence the classification through securing the data until it reaches BS. BS make use of FDA for achieving secure data transfer. Initially, deploy the sensor nodes into the farm field. Respective sensor nodes gives the respective data readings. The data that are sensed at the sensor nodes are sent to the BS. BS apply machine learning techniques to classify the data as anomalous or non-anomalous. Among all the considered supervised learning models, Decision tree works better and give higher percentage of successful classification of the data. The data that belongs to non-anomalous, are sent to second phase of classification ie FDA. The changes in data sensed at sensor nodes at regular interval of time is not so great. Hence it is easy to find the faults in the dataset. BS constructs the matrix for each parameters, by calculating the difference between the received data of each parameters separately.

Then calculate the summation of all values in each rows. If the values are less than or equal to the threshold, then result is 0, which indicates the data detected are non-anomalous. If the summation value obtained is greater than the threshold the result is 1, which shows the data is anomalous. Such anomalous data is considered for the aggregate and such data are dropped. The data concluded as non-anomalous are processed further. The final aggregated data at BS decides to release the water to farm field or not. This notification will be sent to farmers through internet. Famers will take action on irrigation system.

2.1. Data Preparation

The dataset considered for the experimentation is of 3500 records. The parameters or attributes considered are temperature, humidity and soil moisture. For experimentation, the original dataset is divided into three separate dataset of each having different anomaly ratio rates. The target result values are classified in such a way that for anomalous activity, the target value is 1 whereas for non-anomalous activity the target value is 0.

Dataset 1: We added 10% of anomalous records to the dataset.

Dataset 2: 20% of anomalous data are inserted into the dataset.

Dataset 3: Total 30% of anomalous records are included to the dataset.

For each of the dataset considered, the three parameters or attributes with one target value is highlights. Since the data obtained from real world it need to be preprocessed, because it may consist of some missing values or incorrect values. The dataset is split into 70:30 ratio ie 70% of the data are used for training the model whereas 30% of data are used for testing. The sample dataset is represented in Fig 3.

Phase 1: Machine learning techniques for the classification of data as anomalous or non-anomalous.

Machine learning models allows the machine or system to learn without including the programs. In the earlier the system used to give the results automatically, whereas by using the machine learning techniques, the system works like the human brain. The system gives more accurate results. In the smart irrigation scenario, machine learning models or techniques are used to maintain the water resource from wastage of water [7]. Farmers used to make decision based on the traditional practice. Since machine learning models helps the system to work smarter [8][9]. Machine learning techniques used for the classification are K-NN, Multi-Layer Perceptron, Naïve Bayes, Decision Tree, Support Vector Machine, and Logistic Regression. The sensed data are sent to this models. Adding false data into the dataset forbidden results in affecting the crops directly in the farm field by releasing over water or blocking the water supply when it is necessary. Hence it is necessary to rectify the false data in the dataset. The data that are belonging to non-anomalous category, BS apply FDA for checking again. The data that belongs to non-anomalous are sent to further processing to decide the water supply to the farm field.

	Temperature	Humidity	Soil_Moisture	Target
0	28.5826	34.00	0.32004	0
1	28.4930	34.58	0.31990	0
2	28.7616	34.70	0.32060	0
3	28.1762	34.88	0.31990	0
4	27.9586	35.14	0.31990	0
...	...	...	...	...
1647	15.7430	72.16	0.31700	1
1648	16.0198	72.16	0.31700	1
1649	21.6524	72.22	0.29520	1
1650	17.7956	72.28	0.30100	1
1651	16.6918	72.30	0.30100	1

Fig. 3. The sample dataset used for experimentation

Phase 2: Fault Detection Algorithm (FDA)

FDA is invoked for the non-anomalous data. FDA concludes that the data that are received from first phase are non-anomalous, then the data are sent for further processing. BS job is to classify the data. BS apply machine learning techniques and FDA. BS at the second phase creates the matrix M_{i*j}. The matrix created will generate the vector. The majority case at each vector will decide to which category the data belongs to. The algorithm 1 shows in detail the description step by step.

Algorithm: At each sensor nodeS_i, the data D_{S_i} collected at each interval of time t. for each parameter (temperature, humidity, soil moisture) BS creates the matrix M_1, M_2, M_3. The difference between the collected data (D_{S_i} -D_{S_i}). If the difference is less than or equal to threshold (β), then the value is 0 else the value is 1. This is repeated for all the three matrix creation. Summarize the binary values in row-vise, if the value calculated is less the total number of data received divided by 2 then this value decides the data is anomalous or not. Once the vector is obtained if the value is less than the data segment divided by 2, it results in 0 ie, non-anomalous, if it is greater than N/2 then it is 1 ie, the data is anomalous. The vector is created if the value is 0 that shows the data is non-anomalous if data is 1 then the data is anomalous.

The matrix M_1, M_2, M_3 are created according to (1):

$$M_{i*j} = \begin{cases} 0, if \ \left(D_{S_i} - D_{S_j}\right) \leq \beta \\ 1, otherwise \end{cases} \quad (1)$$

Where $m,n = \{0,1,2,\dots x-1\}$, i,j ={0,1,2,…N}, β is threshold

For each row in matrix M_1, M_2, M_3, R_1, R_2, R_3 are calculated as:

$$R_i = \begin{cases} 0, if \ \left(\sum_{j=t-x+1}^{n} M_{i*j} < N/2\right) \\ 1, otherwise \end{cases} \quad (2)$$

At time t, the value of R_i is corrected by R_i^t.

$$R_i^t \begin{cases} 0, if \ \left(\sum_{r=t-x+1}^{t} R_i < {}^{N}/_{2}\right) \\ 1, otherwise \end{cases} \quad (3)$$

For example, let us assume that number of data collected at cluster leader N = 5 and β is 2. The values received from the sensors are temperature= {22.3, 25.2, 23.6, 28.2, 29.3} humidity= {20.8, 22.3, 20.9, 29.3, 20.2} and soil moisture = {50.1, 53.2, 52.2, 63.6, 54.9}. Then the matrices are obtained based on (1) as given in M1, M2 and M3.

$$M_1 = \begin{bmatrix} 0&0&0&1&1 \\ 0&0&0&0&1 \\ 0&0&0&1&1 \\ 1&0&1&0&0 \\ 1&1&1&0&0 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0&0&0&1&0 \\ 0&0&0&1&0 \\ 0&0&0&1&0 \\ 1&1&1&0&1 \\ 0&0&0&1&0 \end{bmatrix} \quad M_3 = \begin{bmatrix} 0&0&0&1&1 \\ 0&0&0&1&0 \\ 0&0&0&1&0 \\ 1&1&1&0&1 \\ 1&0&0&1&0 \end{bmatrix}$$

Further, add the values of each row in a matrix according to equation (2) to obtain vectors R1, R2 and R3 respectively for M1, M2 and M3.

$$R_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad R_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad R_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Take the majority value to decide the node as anomalous or not. The value received when $R_i = 0$ is considered as it is at time t.

Algorithm: The pseudocode for fault detection algorithm

BEGIN

For *every node S_i in WSN (1, 2,…N)*

/ the method helps to create the matrix*/*

If $\left(D_{S_i} - D_{S_j}\right) \leq \beta$

M_{i*j}=0

Else

M_{i*j} =1

End If

End For

*/*construct vector R */*

For *i,j =1,2,…,N*

If $R_i = \sum_{j=t-x+1}^{n} M_{i*j} < N/2$

R_i = 0

Else

R_i =1

End If

If R_i^t = $\left(\sum_{r=t-x+1}^{t} R_i < {}^{x}/_{2}\right)$

R_i^t =0

Else

R_i^t=1

End If

End For

End

Take the majority number to decide the data as legitimate or not. The detailed explanation of FDA is shown in algorithm 1. The previous data that are sensed by sensors are stored to classify the data.

Table 1 shows data ranges in accordance with the parameter which are used to make the decision regarding the water supply. The data that are obtained from the previous round are stored that are used for further round.

Table 1: the data ranges for each parameter

Temperature (°C)	Humidity (%)	Soil Moisture (%)	Decision
>45	<30	<30	Most Required
35-45	35-45	30-45	Required
25-35	46-60	46-60	Average
20-24	61-80	61-80	Not Required
<20	>80	80-100	Not at all required

3. Results and Discussion

The predictions are made in this section. The quality of prediction helps to decide better. The prediction are to classify the data as anomalous or non-anomalous. The metrics obtained while classification of data are temperature, humidity and soil moisture content. This will predict anomalous data whose target value is 1 whereas for non-anomalous data the target value is 0. The three scenarios are recorded here,

Dataset 1:

Total number of instances received after pre-processing the data: 1652

Total number of instances considered for testing: 496

Total number of Normal Activity: 1496

Total number of Anomaly Activity: 156

The table 2 gives classification report of Dataset 1

The classification report exhibits the precision, Recall, F1 Score, support of the models. The same procedure is repeated for all the three datasets. The same algorithms are used for the classification of data

Table 2: The classification report for Dataset 1

Algorithm	Input	**Precision**	**Recall**	**F1 score**	**Support**
Decision Tree	0	0.96	0.95	0.95	453
	1	0.51	0.53	0.52	43
	Weighted Average	0.92	0.92	0.92	496
K-Nearest Neighbor	0	0.93	0.98	0.96	453
	1	0.55	0.28	0.37	43
	Weighted Average	0.90	0.92	0.92	496
Support Vector Machine	0	0.91	1.00	0.92	453
	1	0.00	0.00	0.00	43
	Weighted Average	0.83	0.91	0.87	496
Multilayer Perceptron	0	0.91	1.00	0.95	453
	1	0.00	0.00	0.00	43
	Weighted Average	0.83	0.91	0.87	496
Logistic Regression	0	0.91	1.00	0.95	453
	1	0.00	0.00	0.00	43
	Weighted Average	0.83	0.91	0.87	496
Naïve Bayes	0	0.96	0.96	0.96	453
	1	0.55	0.53	0.54	43
	Weighted Average	0.92	0.92	0.92	496

Table 3: The classification report of dataset 2

Algorithm		**Precision**	**Recall**	**F1 score**	**Support**
Decision Tree	0	0.94	0.93	0.94	450
	1	0.67	0.70	0.68	90
	Weighted Average	0.89	0.89	0.89	540
K-Nearest Neighbor	0	0.88	0.97	0.92	450

	1	0.70	0.34	0.46	90
	Weighted Average	0.85	0.87	0.85	540
Support Vector Machine	0	0.83	1.00	0.91	450
	1	0.00	0.00	0.00	90
	Weighted Average	0.69	0.83	0.76	540
Multilayer Perceptron	0	0.83	1.00	0.91	450
	1	0.00	0.00	0.00	90
	Weighted Average	0.69	0.83	0.76	540
Logistic Regression	0	0.83	1.00	0.91	450
	1	0.00	0.00	0.00	90
	Weighted Average	0.69	0.83	0.76	540
Naïve Bayes	0	0.94	0.93	0.94	450
	1	0.67	0.70	0.68	90
	Weighted Average	0.89	0.89	0.89	540

Table 4: The classification report for Dataset 3

Algorithm		**Precision**	**Recall**	**F1 score**	**Support**
Decision Tree	0	0.94	0.93	0.94	450
	1	0.67	0.70	0.68	90
	Weighted Average	0.89	0.89	0.89	540
K-Nearest Neighbor	0	0.88	0.97	0.92	450
	1	0.70	0.34	0.46	90
	Weighted Average	0.85	0.87	0.85	540
Support Vector Machine	0	0.83	1.00	0.91	450
	1	0.00	0.00	0.00	90
	Weighted Average	0.69	0.83	0.76	540
Multilayer Perceptron	0	0.83	1.00	0.91	450
	1	0.00	0.00	0.00	90
	Weighted Average	0.69	0.83	0.76	540
Logistic Regression	0	0.83	1.00	0.91	450
	1	0.00	0.00	0.00	90
	Weighted Average	0.69	0.83	0.76	540
Naïve Bayes	0	0.94	0.93	0.94	450
	1	0.67	0.70	0.68	90
	Weighted Average	0.89	0.89	0.89	540

Table 5: The confusion metrics and accuracy for each datasets

ALGORITHM	CONFUSION MATRIX			ACCURACY		
	DataSet1	**DataSet2**	**DataSet3**	**DataSet1**	**DataSet2**	**DataSet3**
Decision Tree	$CM = \begin{bmatrix} 417 & 29 \\ 25 & 130 \end{bmatrix}$	$CM = \begin{bmatrix} 419 & 31 \\ 27 & 63 \end{bmatrix}$	$CM = \begin{bmatrix} 431 & 22 \\ 20 & 23 \end{bmatrix}$	91.01%	91.01%	91.53%
K Nearest Neighbor	$CM = \begin{bmatrix} 428 & 18 \\ 56 & 99 \end{bmatrix}$	$CM = \begin{bmatrix} 437 & 13 \\ 59 & 31 \end{bmatrix}$	$CM = \begin{bmatrix} 443 & 10 \\ 31 & 12 \end{bmatrix}$	87.68%	87.68%	91.73%
Support Vector Machine	$CM = \begin{bmatrix} 434 & 12 \\ 70 & 85 \end{bmatrix}$	$CM = \begin{bmatrix} 450 & 00 \\ 90 & 00 \end{bmatrix}$	$CM = \begin{bmatrix} 453 & 00 \\ 43 & 00 \end{bmatrix}$	86.35%	86.35%	91.33%
Multilayer Perceptron	$CM = \begin{bmatrix} 434 & 12 \\ 70 & 85 \end{bmatrix}$	$CM = \begin{bmatrix} 450 & 00 \\ 90 & 00 \end{bmatrix}$	$CM = \begin{bmatrix} 453 & 00 \\ 43 & 00 \end{bmatrix}$	86.35%	86.35%	91.33%
Logistic Regression	$CM = \begin{bmatrix} 434 & 12 \\ 70 & 85 \end{bmatrix}$	$CM = \begin{bmatrix} 450 & 00 \\ 90 & 00 \end{bmatrix}$	$CM = \begin{bmatrix} 453 & 00 \\ 00 & 00 \end{bmatrix}$	86.35%	86.35%	91.33%
Naïve Bayes	$CM = \begin{bmatrix} 417 & 29 \\ 25 & 130 \end{bmatrix}$	$CM = \begin{bmatrix} 419 & 31 \\ 27 & 63 \end{bmatrix}$	$CM = \begin{bmatrix} 434 & 19 \\ 20 & 23 \end{bmatrix}$	73.54%	73.54%	79.43%

Dataset 2:

Total number of instances received after pre-processing the data: 1798

Total number of instances considered for testing: 540

Total number of Normal Activity: 1496

Total number of an Anomaly Activity: 302

The Table 3 gives classification report for Dataset 2.

Dataset 3:

Total number of instances received after pre-processing the data: 2002

Total number of instances considered for testing: 601

Total number of Normal Activity: 1496

Total number of an Anomaly Activity: 506

The Table 4 gives classification report for Dataset3

Accuracy is calculated for the machine learning techniques

Using the following formula,

$$Accuracy = \frac{True\ positive + True\ Negative}{True\ positive + False\ positive + True\ Negative + False\ Negative}$$

Table 5 gives the report of confusion metrics and Accuracy for all the datasets.

By looking into the classification reports of each tables by applying all the machine learning models, we plotted the comparison graph in figure 4.

Apart from the background techniques in the scheme, the front end side, we need to enter five instances. If all the instances entered are in the range then we get the pop up message as non-anomalous data. Figure 5 shows the screenshot of the webpage. The screenshot enlighten after applying both the classification phases to the data records.

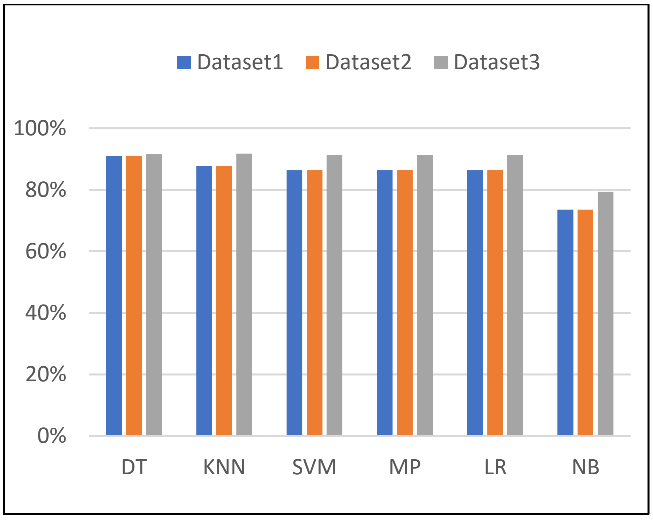

Fig. 4: Comparison of Algorithms

Figure 6 shows the screenshot of the webpage that display anomalous message when one or more instances entered fall out of range. By this way we can conclude that FDA works better for classifying the data.

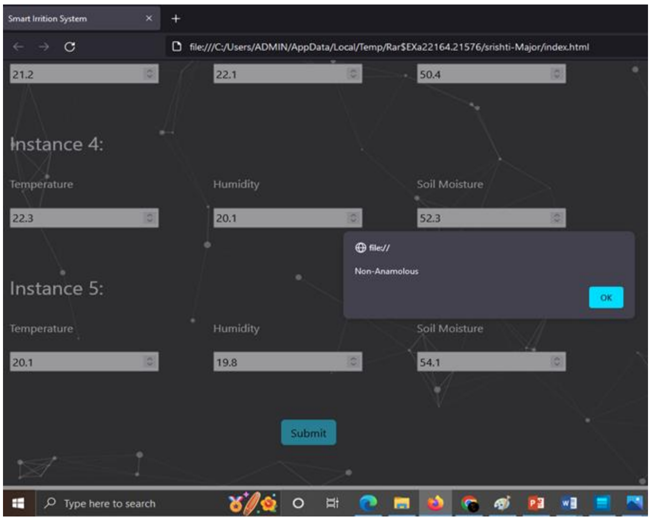

Fig. 5: The screenshot of the webpage

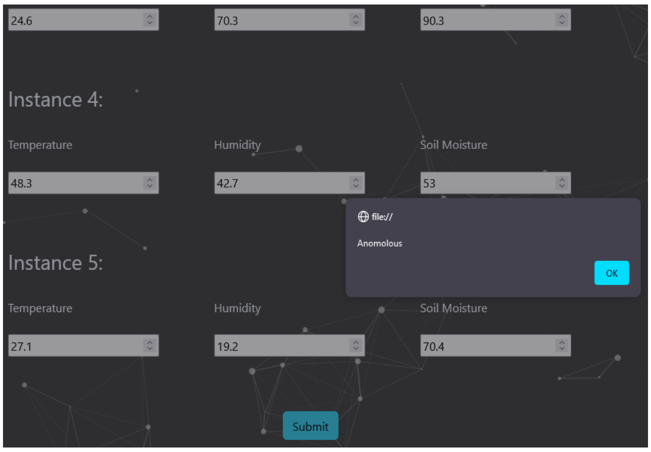

Fig. 6: The screenshot of the webpage displaying message as Anomalous

4. Conclusion

The basic occupation of the world is farming. Hence it is necessary to monitor regularly. This scheme helps the farmer to classify the anomalous and non-anomalous data before making a decision regarding water supply. In the first phase, the data classified into anomalous or non-anomalous data using machine learning techniques. Only non-anomalous data is forwarded to second phase, which is FDA. Once the data is confirmed as non-anomalous, then it is considered for decision making. Based on this data collected from sensors, the farmers take the decision to allow water supply or deny water supply. This scheme provides security for the irrigation system.

References

[1] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, "Machine learning in agriculture: A review," *Sensors (Switzerland)*, vol. 18, no. 8, pp. 1–29, 2018, doi: 10.3390/s18082674.

[2] K. Jha, A. Doshi, P. Patel, and M. Shah, "A comprehensive review on automation in agriculture using artificial intelligence," *Artif. Intell. Agric.*, vol. 2, pp. 1–12, 2019, doi: 10.1016/j.aiia.2019.05.004.

[3] N. Khan, R. L. Ray, G. R. Sargani, M. Ihtisham, M. Khayyam, and S. Ismail, "Current progress and future prospects of agriculture technology: Gateway to sustainable agriculture," *Sustain.*, vol. 13, no. 9, pp. 1–31, 2021, doi: 10.3390/su13094883.

[4] T. Ojha, S. Misra, and N. S. Raghuwanshi, "Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges," *Comput. Electron. Agric.*, vol. 118, pp. 66–84, 2015, doi: 10.1016/j.compag.2015.08.011.

[5] N. G. Rezk, E. E. D. Hemdan, A. F. Attia, A. El-Sayed, and M. A. El-Rashidy, "An efficient IoT based smart farming system using machine learning algorithms," *Multimed. Tools Appl.*, vol. 80, no. 1, pp. 773–797, 2021, doi: 10.1007/s11042-020-09740-6.

[6] T. K. R. S Akshay, "Efficient Machine Learning Algorithm for Smart Irrigation," 2020, [Online]. Available: https://www.researchgate.net/publication/344063809_Efficient_Machine_Learning_Algorithm_for_Smart_Irrigation.

[7] S. K. S. Durai and M. D. Shamili, "Smart farming using Machine Learning and Deep Learning techniques," *Decis. Anal. J.*, vol. 3, no. December 2021, p. 100041, 2022, doi: 10.1016/j.dajour.2022.100041.

[8] S. Ramya, A. M. Swetha, and M. Doraipandian, "IoT framework for smart irrigation using machine learning technique," *J. Comput. Sci.*, vol. 16, no. 3, pp. 355–363, 2020, doi: 10.3844/JCSSP.2020.355.363.

[9] A. Senthilkumar, "Minimizing Energy Consumption in Vehicular Sensor Networks Using Relentless Particle Swarm Optimization Routing," vol. 10, no. 2, pp. 217–230, 2023, doi: 10.22247/ijcna/2023/220737.

[10] M. Pathan, N. Patel, H. Yagnik, and M. Shah, "Artificial cognition for applications in smart agriculture: A comprehensive review," *Artif. Intell. Agric.*, vol. 4, pp. 81–95, 2020, doi: 10.1016/j.aiia.2020.06.001.

[11] E. T. Sigfredo Fuentes, "Advanced and Requirements for machine learning and artificial intelligence applications in viticulture," 2018, [Online]. Available: https://www.researchgate.net/publication/325710884_Advances_and_requirements_for_machine_learning_and_artificial_intelligence_applications_in_viticulture.

[12] Y. Mekonnen, S. Namuduri, L. Burton, A. Sarwat,

and S. Bhansali, “Review—Machine Learning Techniques in Wireless Sensor Network Based Precision Agriculture,” *J. Electrochem. Soc.*, vol. 167, no. 3, p. 037522, 2020, doi: 10.1149/2.0222003jes.

[13] Y. Akshatha and A. S. Poornima, “IoT Enabled Smart Farming: A Review,” *Proc. - 2022 6th Int. Conf. Intell. Comput. Control Syst. ICICCS 2022*, no. Iciccs, pp. 431–436, 2022, doi: 10.1109/ICICCS53718.2022.9788149.

[14] M. Sahu, N. Sethi, S. K. Das, and U. Ghugar, “NDTRA-MAT: A Novel Technique for Evaluating the Data Transfer Rate, Reducing the False Alarm Rate, and avoiding Packet Droppings Rate against Malicious Activity in Wireless Sensor Networks,” *Int. J. Comput. Networks Appl.*, vol. 10, no. 1, p. 1, 2023, doi: 10.22247/ijcna/2023/218507.

[15] P. Rutravigneshwaran and G. Anitha, “Security Model to Mitigate Black Hole Attack on Internet of Battlefield Things (IoBT) Using Trust and K-Means Clustering Algorithm,” *Int. J. Comput. Networks Appl.*, vol. 10, no. 1, p. 95, 2023, doi: 10.22247/ijcna/2023/218514.

[16] T. S. Othman and S. M. Abdullah, “Intrusion Detection Systems for IoT Attack Detection and Identification Using Intelligent Techniques,” *Int. J. Comput. Networks Appl.*, vol. 10, no. 1, p. 130, 2023, doi: 10.22247/ijcna/2023/218517.

[17] T. M. Alshammari and F. M. Alserhani, “Scalable and Robust Intrusion Detection System to Secure the IoT Environments using Software Defined Networks (SDN) Enabled Architecture,” *Int. J. Comput. Networks Appl.*, vol. 9, no. 6, pp. 678–688, 2022, doi: 10.22247/ijcna/2022/217701.

[18] S. Venkatasubramanian, A. Suhasini, and S. Hariprasath, “Detection of Black and Grey Hole Attacks Using Hybrid Cat with PSO-Based Deep Learning Algorithm in MANET,” *Int. J. Comput. Networks Appl.*, vol. 9, no. 6, pp. 724–735, 2022, doi: 10.22247/ijcna/2022/217705.

[19] Y. Akshatha, A. S. Poornima, and M. B. Nirmala, “Secure Data Collection in Clustered Wireless Sensor Networks using Fuzzy based scheme to detect Malicious Data Collector,” *Int. J. Eng. Trends Technol.*, vol. 70, no. 11, pp. 240–248, 2022, doi: 10.14445/22315381/IJETT-V70I11P226.

[20] A. Angelaki, S. Singh Nain, V. Singh, and P. Sihag, “Estimation of models for cumulative infiltration of soil using machine learning methods,” *ISH J. Hydraul. Eng.*, vol. 27, no. 2, pp. 162–169, 2021, doi: 10.1080/09715010.2018.1531274.

[21] C. Crisci, B. Ghattas, and G. Perera, “A review of supervised machine learning algorithms and their applications to ecological data,” *Ecol. Modell.*, vol. 240, no. June 2018, pp. 113–122, 2012, doi: 10.1016/j.ecolmodel.2012.03.001.

[22] M. Çetin, S. Yıldız, and S. Beyhan, “Water Need Models and Irrigation Decision Systems,” no. July, 2021, [Online]. Available: http://arxiv.org/abs/2103.11133.

[23] T. Shakoor, “Intelligent Agricultural Information Monitoring Using Data Mining Techniques,” 2017.

[24] M. S. Munir, I. S. Bajwa, A. Ashraf, W. Anwar, and R. Rashid, “Intelligent and Smart Irrigation System Using Edge Computing and IoT,” *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/6691571.

[25] A. Vangala, A. K. Das, V. Chamola, V. Korotaev, and J. J. P. C. Rodrigues, “Security in IoT-enabled smart agriculture: architecture, security solutions and challenges,” *Cluster Comput.*, vol. 26, no. 2, pp. 879–902, 2022, doi: 10.1007/s10586-022-03566-7.

[26] A. Vangala, A. K. Das, V. Chamola, V. Korotaev, and J. J. P. C. Rodrigues, “Security in IoT-enabled smart agriculture: architecture, security solutions and challenges,” *Cluster Comput.*, no. April, 2022, doi: 10.1007/s10586-022-03566-7.

[27] S. Raveena and A. Shirly Edward, “Secure B-IoT Based Smart Agriculture- A Brief Review,” *J. Phys. Conf. Ser.*, vol. 1964, no. 4, 2021, doi: 10.1088/1742-6596/1964/4/042006.

[28] Naidu k, P. ., Rao, V. L. ., Gunturu, C. S. ., Niharika, A. ., Anupama, C. R. ., & Srivalli, G. . (2023). Crop Yield Prediction Using Gradient Boosting Neural Network Regression Model . International Journal on Recent and Innovation Trends in Computing and Communication, 11(3), 206–214. https://doi.org/10.17762/ijritcc.v11i3.6338

[29] Mr. Dharmesh Dhabliya. (2012). Intelligent Banal type INS based Wassily chair (INSW). International Journal of New Practices in Management and Engineering, 1(01), 01 - 08. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/2