2022 5th International Conference on Contemporary Computing and Informatics (IC3I)

IoT and ML based Irrigation System using KNN Algorithm

Dr. Arvind Kumar Pandey
Assistant Professor
Department of Computer Science
ARKA JAIN University
Jamshedpur, Jharkhand, India
dr.arvind@arkajainuniversity.ac.in

Prof. Yoothika Patel
Assistant Professor
Department of CSE
PIET, Parul University
Vadodara, Gujarat, India.
yoothika.patel270041@paruluniversity.ac.in

***Abstract*—Agriculture is an important sector that contributes significantly to the improvement of the social and economic conditions of every nation around the globe. Water is essential to life and also represents a valuable resource for agricultural production. Emerging trends in agriculture have the goals of maximizing crop yields while minimizing harm to the environment and contributing to the development of sustainable agricultural practices (12). The level of moisture in the soil is the single most important factor in determining crop productivity; insufficient moisture leads to a reduction in yield and ultimately the death of the plants. Wasted water and illnesses of the roots can be caused by an abundance of water. It is necessary to apply the appropriate amount of water at the optimal moment in order to achieve maximum crop output. The IOT and ML will be used to construct an automated watering system in the proposed project. The system is broken down into three stages: (1) building the circuit and installing the sensors, (2) involves observing the sensor data and determining the current level of soil moisture, (3) involves determining whether or not to turn the motor on or off based on the type of crop being grown. The data from the sensors is saved on the server, and the system provides data from the sensors in real time. The IOT (IoT) assists in the implementation of Wireless Sensor Networks, and ML can manage the most efficient use of scarce water resources for farming by analyzing the data that is readily available.**

Keywords*—Support Vector Machine (SVM), Machine learning (ML), Agriculture, K-Nearest Neighbor (KNN),Irrigation, Decision Tree (DT), Internet of Things (IOT), Microcontroller.

I. Introduction

The technique of irrigating agricultural land helps to limit the amount of water that is wasted while also ensuring that water is distributed effectively across the area. Farmers are forced to contend with irregular rainfall, which has a negative impact on food production and, as a result, will impede the socioeconomic status of farmers. The rising number of farmers who take their own lives is one indication of this trend. This is also attributed, in part, to the development of more efficient farming, irrigation, and water-source practises. Therefore, their socioeconomic condition is one of the worst in the world. As a result, there is a need for technology that can anticipate how much water a crop will require based on the soil it will grow in.

A wireless sensor-based implementation combined with ML helps to boost output while using a sufficient amount of water. This helps increase productivity on farms, as well as profitability and efficiency. The traditional method of irrigation results in significant losses of both water and electricity. Ineffective management of irrigation can lead to problems such as over- or under-irrigation, both of which can reduce crop output while also causing damage to crops, a reduction in nutrient uptake, and a drop in aeration. The way in which irrigation affects the productivity of land is through the role it plays in determining the level of cultivation (number of cropping seasons). [13] In order to keep the same level of soil moisture an enhanced irrigation, which means an automated water delivery system for the plants is being implemented. Taking into account the current temperature, humidity, and soil moisture, the motor for the water pump can be programmed to switch on and off automatically. The purpose of the opted study is to build automatic irrigation systems &attempt to make more informed decisions regarding them. The work of surveillance and managing the automatic irrigation process is carried out by this system. [14]

A. An enhanced water distribution system based on IoT

IoT is a difficult topic of computer science; it refers to the intelligent collection and processing of data. Automation of scientific implication can be made reliable by using sensors and microcontrollers in order to supply the required amount of water at the measured quantity at the required time. The data from the sensor can be fed into the microcontroller in some fashion. The microcontroller has the ability to regulate and monitor the irrigation process. Devices are able to connect with one another through the core Network in an IOT (IoT)-based system. The sensor is used to collect data. It is possible to use technologies such as Arduino Uno and Raspberry Pi3 for the development and communication of the devices.

B. Utilizing ML to Create an Enhanced Irrigation System

Artificial intelligence and computer science have merged to form the discipline known as machine learning. It offers a data analysis method and model that may be created by employing a method to learn information from the data. The process of a aiming the irrigation procedure with a variety of tools and a prediction model can benefit from machine

979-8-3503-9826-7/22/$31.00 ©2022 IEEE

2022 5th International Conference on Contemporary Computing and Informatics (IC3I) | 979-8-3503-9826-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/IC3I56241.2022.10072613

learning. It is possible to implement it in an automated irrigation system by taking into consideration the moisture content of the field's soil. The real-time parameters help to give accurate decisions that contribute to an efficient agricultural process. (17).

II. Literature Review

In agriculture, water management can be approached from a variety of contemporary perspectives, including the following:

The author stated that irrigation is the most important factor in agriculture since the lack of adequate irrigation is the most significant barrier to expanding agricultural crop production.

The author makes a number of recommendations to be implemented for major benefits; to properly manage water, one should think globally but act locally; funds from the cooperative sector need to be raised at the district level; and so on. [1]

Using Arduino, microcontroller ESP8266, and certain fundamental sensors, created and constructed a mobile application with the purpose of automatically monitoring and regulating an irrigation system. The author intends to make his system better by incorporating a user interface with a facility for setting a range of acceptable soil moisture levels. In addition, the functionality of the system may be enhanced by incorporating some form of intelligence that would analyse the dataset and instruct the computer to make correct decisions. [2]

ML offers a variety of applications that may be used to analyse data in a more complex manner. This aids in the taking of decisions that are more correct, which increases agricultural productivity. The authors highlight the ways in which ML can be applied in the field of agriculture. [3]

A system for automatically watering plants and identifying fires has been developed through the use of a variety of sensors and cloud computing. The use of these technologies contributes to the advancement of agriculture. [4]

The author presented a method that decentralizes monitoring and decision making in the field of agricultural irrigation. The microcontroller, the sensors, and the integration of the water pumps with the decision-making system are the components that make up the system. [5]

ML was applied to improve the system's accuracy. At the beginning of the process, the amount of water needed for irrigation is determined with the assistance of a advanced irrigation system and data from the surrounding environment. [6]

Using a microcontroller, sensors, and Bluetooth created an data collection method makes use of a variety of sensors, and the results are then compared to specified thresholds. The Arduino Board that was utilized in the process of connecting the sensors. The amount of moisture in the soil determines whether or not the pump is turned on or off [7].

This study [8] provides an illustration of the categorization challenge applicable to automated irrigation processes. The author claim that sensor-based data is gathered from the area and the weather forecast in order to predict soil moisture using ML techniques. A number of alternative ML methods, including Gradient Boosting Regression (GBRT), Multiple Linear Regression (MLR), Elastic Net Regression (ENR), and Random Forest Regression, are being investigated for their ability to forecast soil moisture (RFR). After looking at all of these different methods, the GBRT has proven to produce the most promising results for predicting soil moisture [9].

III. Proposed System

Figure 1 illustrates the development architectural model that will be used for the proposed study. In the process of irrigating crops, the moisture content of the soil is the most crucial characteristic to consider. Inadequate or excessive water can cause plant diseases and lead to water waste. In a similar vein, the current temperature and humidity of the surrounding environment helps to make an informed selection regarding the watering process. The investigation is broken up into a total of six parts (Figure 2).

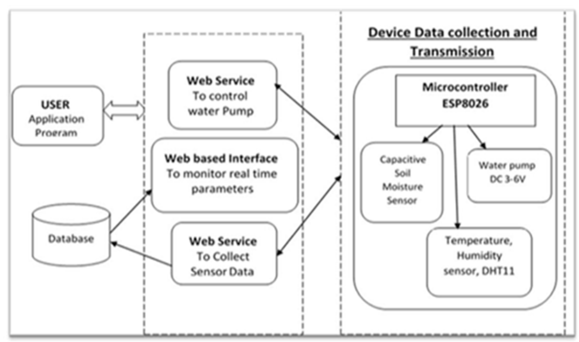

Figure 1. Project Architectural model

Figure 2. Schematic architectural.

A. The Installation of Sensors and the Establishment of Circuits

Sensors, a relay switch, a web service, and a water motor were used to construct the circuit. The sensors are installed in the correct locations and then jumper wires are used to connect them to the ESP8266 microcontroller shown in Figure 3. PHP is the programming language that is used to construct the web-based interface that allows users to examine real-time data quantities. The user interface allows for the setting of both a minimum and maximum amount of soil moisture. The process of turning on and off the motor has been automated while taking into account the soil's moisture content at the time. The water motor will turn on if the current moisture in the soil level is below the planned minimum soil moisture level. The pump motor will shut off when the soil moisture level exceeds the predetermined maximum. The water motor won't turn on if the current soil moisture level exceeds the planned minimum soil moisture level.

A link to the internet via Wi-Fi is necessary for the processes of data collecting and monitoring, as well as for the automatic irrigation system. PHP and MySQL are used in the development of the user interface. Using this web-based interface, you can monitor real-time data as well as the state of the water motor in real time.

In order to turn the water motor on and off, it is connected to a relay switch as well as a web service. The user interface for the water motor allows for both manual and automatic operation of the device.

B. Temperature and Humidity sensor known as the DHT11

This environmental humidity and temperature sensor is of high quality and very reasonable pricing. In addition, there is a low consumption of power, which is stable over the long period. Temperatures can range anywhere from 0 to 50 degrees Celsius, while humidity can be anywhere from 20% to 80%.

C. V2.0 Soil Moisture Sensor

Utilizing the top of the plant root, this transducer measures the soil's moisture content. Here, the capacitive soil humidity transducer is utilized, which uses capacitive sensing to keep track of the moisture content of the soil. We selected a capacitive moisture sensor despite the fact that there are other varieties of soil moisture sensors on the market. This is a very reliable moisture sensor that also comes at a very reasonable price. This sensor can measure from 0 all the way up to 1023.

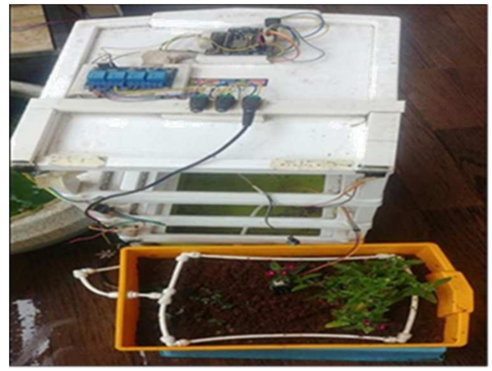

Figure 3. Experimental Setup

IV. Data Collection

The microcontroller ESP8266 read the current data from the temperature and humidity sensor, as well as the soil moisture sensor. In order to collect data at regular 15-minute intervals, the necessary programming is done in Micro Python. The server receives sensor data that is sent from the ESP microcontroller. The values for moisture, temperature, and humidity of the soil are recorded together with the date and time for a period of twenty days. Therefore, there are a total of 1920 observations that may be analysed in this study. The model is wired to an uninterruptible power supply (UPS) backup system in order to capture continuous data at 15-minute intervals.

A. The Processing of Data and the Extraction of Features

This is a very crucial phase in the process of developing a ML model, as it involves cleaning and organizing the data in an appropriate form. This helps to increase the accuracy and quality of the ML Model.

B. ML model development

In the study that was proposed, three different supervised classification algorithms were applied to the dataset that was available. We investigate the KNN, ML classifier, the DT classifier, and the SVM algorithm, and we acquire findings that compare accuracy, confusion matrix, and classification matrix.

During the process of data mining, enormous datasets are sorted. After that, pattern recognition must be carried out, and then one must carry out performance analysis in order to develop a relationship and solve the problem. The act of distinguishing between the various data classes and concepts and describing them in preparation for the process of model finding is referred to as the data classification process. The dataset was divided into two parts: the training dataset and the testing dataset. The training dataset contains 80% of the total data so that a machine classifier may be trained on it and a ML method can be developed using it. The testing dataset includes twenty percent of the total data in order to test the model.

C. KNN

The KNN algorithm searches for data points nearby that are comparable to the ones it's looking for, and then finds the output. No training is required for the KNN method, which only takes into account the minimal distance between neighbours. [15]

D. The Tree of Decisions

The construction of the prediction model is carried out using a method based on trees. It is obtained by the use of an independent variable, where each node has a condition over a feature. The condition will determine which node will be navigated to after this one. When the output was supposed to have been reached by the leaf node. It is one of the simpler algorithms, and it is easier to grasp and put into practice. Additionally, it produces more accurate results. This method is utilized to provide assistance with decision making and risk assessment. [11].

E. SVM

When attempting to discover non-linear boundaries and linear boundary than modified version of feature will be beneficial. It represents training data as dots in space, with significant distances separating each category. [16].

F. Practice and Examination of the Model

After the stage when the data is preprocessed, the next step in constructing the ML model is to divide the dataset into the training dataset and the testing dataset. The dataset is typically divided into a training dataset and a testing dataset in order to facilitate the process of fitting a ML model and estimating the performance of this model. This model is constructed with the assistance of various ML algorithms in order to predict certain results. In this step, we fit our model to an existing dataset in which both the inputs and outputs are known. After that, we prepare predictions based on data that is either unknown or will be input in the future. This technique of training and testing can only operate appropriately when a sufficiently large dataset is provided. The test data set was applied to the training dataset in order to conduct an objective evaluation of the model that was generated. In the current study, the actual dataset is divided into an 80-20 pattern, which indicates that the training dataset size accounts for 80% of the total, and the testing dataset accounts for the remaining 20%. We receive a total of 1388 observations in the training data from the dataset that is currently available, and we get a total of 347 observations in the testing dataset.

Using the train test split () function found in the Scikit-learn library, which is part of the Python ML library, carry out the train and test splitting operation.

G. Forecasting the Outputs

The both precision and the efficacy are scrutinized in the area of the evaluation that is devoted to the results. Additionally, the observation of accurate irrigation ON/OFF results is made possible by the supply of input data sets for input parameters. In this section, the outcomes of the algorithms that were put into action are presented in Table I, Table II, AND Table III.

TABLE I. CLASSIFICATION REPORT KNN

	Support	Precision	F1-Score	Recall
WEIGHTED AVG	347	0.98	0.98	0.99
OFF	343	0.99	0.99	1
ACCURACY	347		0.99	
ON	4	0	0	0
MACRO AVG	347	0.49	0.5	0.5
KNN				

TABLE II. CLASSIFICATION REPORT DT

	support	precision	f1-score	recall
weighted avg	347	1	1	1
0	343	1	1	1
accuracy	347		1	
1	4	1	1	1
macro avg	347	1	1	1
DT				

TABLE III. CLASSIFICATION REPORT SVM

	support	precision	f1-score	recall
weighted avg	347	0.99	0.78	0.64
0	343	1	0.78	0.64
accuracy	347		0.64	
1	4	0.36	0.53	1
macro avg	347	0.45	0.44	0.55
8	0	0	0	0

Accuracy can be defined as the degree to which the expected observation corresponds correctly to the overall observation. The degree to which the results that were predicted were accurate when they were determined for the performance calculation of the model. The amount of times one's forecast was accurate can either go up or down, and these extremes can be thought of as true positives and true negatives respectively.

As a metric of relevance, precision and recall are both utilized. Recall refers to the percentage of relevant instances retrieved from the entire number of instances, while precision refers to the function of relevant instances retrieved from total instances.

The F1-Score, also known as F Measure or F Score, is the weighted average score obtained by averaging the precision and recall scores. The formula for determining it is 2 times (precision * recall) divided by (precision plus recall). It gives the impression of precision while also evoking recollection.

V. RESULT

The outcomes of the suggested model's predictions are more accurate than those of any other model and have a lower rate of error. The procedure of irrigating was controlled through a web interface. The model is fully functioning, and the results of the predictions are likewise of a high quality. According to the findings that were predicted by all three algorithms, the accuracy rate of the KNN algorithm is 99%, the accuracy rate of the DT algorithm is 100%, and the accuracy rate of SVM is 64%. The model for the smart autonomous irrigation system is constructed using an algorithm for machine learning. [11] The data is gathered through the use of a variety of sensors, and the DT (DT) algorithm can accurately anticipate the result based on the data with 97.86% precision. Anat Goldstein and colleagues also constructed a model by employing a variety of regression and classification algorithms in their work. The Gradient Boosted Regression Tree regression approach achieves an accuracy of 93%, and the Boosted Tree Classifier classification model has an accuracy of 95%. According to the results that were proposed, the DT provides more accurate results, and it may be more suitable for use in automatic irrigation systems. This model assists in the prediction of the outcomes of the irrigation process in the form of irrigation decisions regarding whether or not to turn the water motor on or off. This method is adaptable for use in the garden of an agricultural field, a greenhouse or polyhouse, the garden of a housing society, or even in the garden of a private residence.

VI. FUTURE GUIDELINE

With the use of numerous nodes and at a low cost, the system will eventually be able to be used in the real world. Furthermore, the system will have the ability to expand by incorporating new characteristics, like an evapotranspiration prediction using real-time metrological data. The proposed research has the potential to be expanded in order to maintain the crop's fertilizer requirements in accordance with the crop's age on the farm. In a similar vein, monitoring the amount of water in a tank can also be maintained to some extent.

VII. CONCLUSION

Making accurate decisions at the right moment based on fundamental factors and environmental data can help to boost crop output and reduce the risk associated with farming. ML and the IOT can provide this assistance. The value of soil moisture is particularly important to consider in order achieving optimal irrigation and crop health. A number of climatic factors, such as temperature, humidity, and wind speed, can all have an impact on the moisture content of the soil. The need for irrigation can be forecasted using data collected from sensors in the recent past in the study that was proposed. With the use of an IOT centered architectural model and precise irrigation ON and OFF selection forecast using a ML algorithm, this paper aimed to provide a better technique of irrigation. The design of this system was created using open-standard technologies, and its total cost of ownership is quite low. The application of technology to the process of irrigation in agriculture helps to make the most efficient use of water and boosts agricultural output.

REFERENCES

[1] Todkari G. U., Impact of Irrigation on Agriculture Productivity in Solapur District of Maharashtra State, ISSN- 0975-3710, E-ISSN 0975-9107, Vol-4, Issue 1.2012, International Journal of Agriculture Science.

[2] Peace Muyambo (2019) 'Android Enabled Smart Irrigation System using internet' e-ISSN 2395-0056, p-ISSN 2395-0072 www.irjet.net International Research Journal of Engineering and Technology

[3] Dr. Santosh Jagtap et. al (2021), 'Towards application of various MLtechniques In agriculture', ISSN 2214-7853, Elsevier BV

[4] Mrs. S. Sangeetha et. al.(2017), Smart irrigation based on cloud, ISSN-2321-5992, IPASI International Journal of Computer Science(IIJCS).

[5] Kizito Masaba et.al (2016), Design and implementation of a Smart Irrigation System for Improved water-Energy efficiency, DOI: 10.1049/cp.2016.1357, ISBN: 978-1-78561-238-1, Publisher: IET.

[6] Yu-chuan chang et. al.(2019), A MLBased smart Irrigation System with LoRa P2P networks, IEEE, ISSN: 2576-8565, Asia-Pasific Network operations and Management symposium(APNOMS).

[7] Ashwini B. V (2018), A Study on Smart Irrigation System UsingIoTforSurveillanceofCrop-field,DOI:10.14419/ijet.v7i4.5.20109,International Jornak of Enginnering and Technology (4.5)370-373

[8] R. N. Kirtan (2018), Smart Irrigation System using ZigbeeTechnology and Machine learning Techniques. DOI:10.1109 / I2C2SW45816.2018.8997121, ISBN: 978-1-5386-9433-6, IEEE.

[9] Gursimran Singh el. at. (2019), MLBased soil moisture for IOT Based smart Irrigation System, DOI: 10.1016/j.compag.2018.09.040, IEEE.

[10] Anat Goldstein et. al. (2017), Applying MLon sensor data for irrigation Recommendations: reviling the agronomist's tacit knowledge, DOI 10.1007/s11119-017-9527-4, Springer.

[11] Kumar K, Anand S, Yadava RL. Advanced DSP Technique to Remove Baseline Noise from ECG Signal. Int J Electron Comput Sci Eng. , , vol. 1, no. 3, pp. 1013-1019, 2012.

[12] Kumar K, Tanya Aggrawal, Vishal Verma, Suraj Singh, Shivendra Singh, Dr. Lokesh Varshney, “Modeling and Simulation of Hybrid System”, IJAST, vol. 29, no. 4s, pp. 2857 -2867, Jun. 2020.

[13] Kumar K, Varshney L, Ambikapathy, Vrinda, Sachin, Prashant , Namya. Soft Computing and IoT based Solar Tracker. International

Journal of Power Electronics and Drive System (IJPEDS). Vol 12, No 3: September 2021. doi.org/10.11591/ijpeds.v12.i3.pp1880-1889.
[14] Lishan wang(2019), Research and Implementation on MLclassifier based on KNN, DOI: 10.1088/1757-899X/677/5/052038, IOP Conf. Series: Material science and Engineering 677, IOP publishing.
[15] Hastie et.al.(2017), The elements of statistical learning: Data mining, inference and prediction, eBOOK ISBN 978-0-387-84858-7, series ISSN 0172-7397, DOI 10.1007/978-0-387-84858-7, hardcover ISBN 978-0-387-84857-0, Springer.
[16] Paul L. G. (2008), Land Use and Soil Resources, ISBN: 978-1-4020-6778-5, Springer Publication.