International Journal of Engineering & Technology, 7 (2.7) (2018) 179-182

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

A prototype model for continuous agriculture field monitoring and assessment

RatnmalaBhimanpallewar[1]*, Manda Rama Narasingarao[2]

[1]PhD Scholar, CSE Department Koneru Lakshmaiah Education Foundation, India
[2]Professor, CSE Department Koneru Lakshmaiah Education Foundation, India
**Corresponding author E-mail :ratnmalab@gmail.com*

Abstract

Indian farmers are totally dependent on agriculture and livestock for satisfying their basic food and economical needs. Maximum farmers are habitual to take crops continuously with traditional ways without checking the current suitability. Government of India has developed centers to train and provide the information to farmers but everyone don't approach to it. To get the increased yield, usually farmers add fertilizers without understanding requirement which may leads to soil degradation. Proposed solution is an automated system which can monitor major parameters required to estimate suitability for cropping. This system can be made available locally to every farmer. Outcome of this monitoring system can be used to identify particular crop suitability, so that suitable crop can be adopted.

Keywords: *Suitability Level; Machine Learning; Iot; Agriculture Parameter; Sensor .*

1. Introduction

Recent technologies like Big data, IoT (Internet of Things), smart systems are used in Agriculture domain also [1]. But those are applicable for un-fragmented lands in developed countries where adequate resources can be provided artificially for crop production [2]. In developing countries continuous growth in population is creating pressure on food supply. In developing countries like India day by day there is increase in land degradation. It has become major threat in crop production. Many reasons are involved in land degradation like excess use of fertilizers without knowing its requirement [3]. Even crops are chosen without knowing its suitability for current scenario, which leads to degradation in quantity and quality of crop production. Government of India has developed KVK centres (Krishi Vigyan Kendra) to monitor parameters effecting directly on crop production. It provides facility to test the samples like water, soil etc. Farmers has to collect their samples there and get the test report after some days. Yet there is no any system available which can made available to farmers wherever they want and whenever they need results quickly.

Along with KVK centres materials for guiding farmers are available like "Krishi Darshani" [4], annual reports etc. But every small/marginal scale farmers can't approach it every time. In some of the developed countries farmers are using automated systems for knowing few agriculture parameters [5] and the field suitability also [6] and feeding the suitable inputs as well like water, fertilizer etc. [7], [8]. Most of these automated systems are suitable for large un-fragmented land and are not affordable for small/marginal scale farmers. Few precision agriculture systems are available for monitoring parameters but those are restricted with one or two parameters. One of the recent developed system in Philippines is for monitoring nutrients like pH, Nitrogen, Potassium, Phosphorus (N, P, K) and suggest fertilizers on its basis [9]. Crop suitability is dependent on not only soil nutrients but also some geographic and environmental parameters [10], [11]. Proposed system is a prototype model based on IoT, responsible for monitoring of maximum field parameters required for assessing crop suitability, it covers maximum variable parameters. Using those values it also assess the particular crop suitability .

2. Material and Methods

2.1 Material Used

Some of the components used to build IoT based monitoring module are listed out here.

2.1.1Arduino Uno microcontroller

The Arduino Uno is a open source microcontroller board provides set of digital and analog input/output pins. It contains everything needed to support this IoT module. Using the board it can simply connect to a computer and power source as well. The recommended range is 5v to 12v for Arduino Uno.

It is compatible with Wi-Fi module used here. Multiple calibrated sensors are directly connected to it for measuring the soil and environmental parameters

Copyright © 2018Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



Fig. 1. Arduino Uno Microcontroller

2.1.2ESP8266-01WiFi module

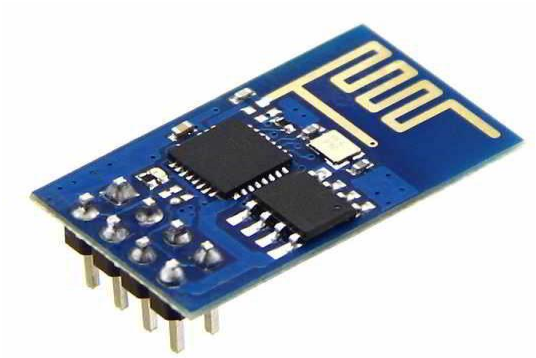


Fig. 2. The ESP8266 WiFi Module

The ESP8266 Wi-Fi Module is a self contained SOC (system on chip) with integrated TCP/IP protocol stack that can give any microcontroller access to your Wi-Fi network. It is capable of either hosting an application or offloading all Wi-Fi networking functions from another application processor. We can simply hook this up to your Arduino device as it is pr-programmed. This cost effective module provides much Wi-Fi-ability as a Wi-Fi Shield offers. It is enough capable for storage and on-board processing. It can be integrated with the sensors and other application specific devices with minimal load. It has self-calibrated RF so that it can work under all operating conditions.

2.1.3Sensors

Some calibrated sensors are used to measure the soil and environmental parameters as listed below.
a) Soil moistures
b) Temperature
c) pH
d) Electrical Conductivity

All above sensors are selected as per their allowed range of values feasible with range of respective parameters in real time agriculture field. As an example circuit diagram connecting to a soil moisture sensor is shown below

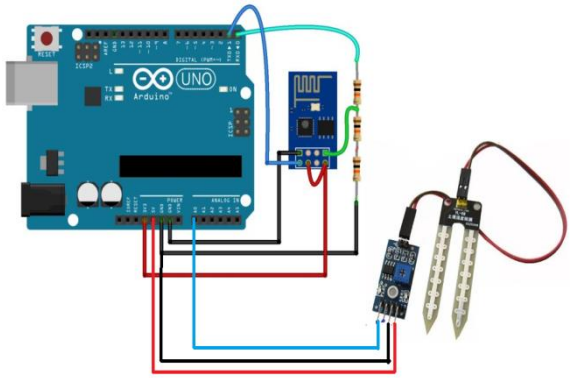


Fig. 3. ESP8266 WiFi interfacing diagram with Arduino Uno and Soil moisture sensor.

Similarly all the sensors are connected appropriately as per their specifications. All the measured parameters are stored locally for specified interval and finally average values will be sent to the cloud. Cloud is managed at server side.

2.2 Method used

Here method used is combination of IoT and hybrid machine learning algorithm. IoT module is used to monitor agriculture variable parameters and other somewhat persistent parameters like Nitrogen, Phosphorus, Potassium (NPK), Soil type, topography are referred from publically available data and reports. Depending on that data, suitability analysis for particular crop is done by a hybrid machine learning algorithm. To train this machine learning model historical data used, which is referred from publically available data from Agriculture University, Pune and published environmental data by Meteorological Department, Pune. Thus scope of current model is Pune region, India which can be extended to scope of any geographical region by feeding the historical data for that particular region to train the model.

2.2.1 Data collection

Multiple nodes are sending data over the cloud. Some constant parameters and variable parameters from nodes are stored using structured database schema. Averaging of the field wise parameters is done at client side as shown in step 2.2.2 then the values are sent over cloud. At server side hybrid machine learning algorithm as in step 2.2.3 is applied for assessing suitability level. Assessment is done only once or twice in agriculture season. Some part of this hybrid machine learning approach published in our previous work also sample historical data is available there [10].

2.2.2 Data preprocessing

let,
Every node is measuring
i=1, 2, ..., n parameters
Every field has
j=1, 2,..., m nodes to monitor all sensor parameters
Xi - any soil/environmental parameter

$$X = \{ X_1, X_2, ..., X_n \} \quad (1)$$

x_{ij} - value of i^{th} parameter at j^{th} node
X_{ij}- average value of i^{th} parameter at j^{th} node i.e. avg (x_{ij})

Let's consider that dataset is belongs to only one field

$$X_1 = \text{average } (x_{1j}) \quad (2)$$

similarly calculate average values of each input parameter.

2.2.3 A Hybrid machine learning algorithm

Input: X - soil and environmental parameters
Output: Suitability level for particular crop depending on quantitative and qualitative approach [12].
S1- Suitable
S2- Moderately suitable
S3- Marginally suitable
N1- Not suitable (due to major economic reasons otherwise moderately suitable)
N2- Not suitable (due to physical reasons).

a) Classification technique c4.5
It is classification technique which can perform multilevel classification. It works with all types of data values categorical and continuous as well [13].

Steps for Algorithm:
I. Divide the dataset T into two parts training dataset and testing dataset. Here output vector is nothing but suitability levels for the cropland,

$$S=\{ S1, S2, S3, N1, N2\} \quad (3)$$

1. X- set of every parameter considered as input, there are n- number of input parameters as we know already
2. Select one parameter P :output(S) or input(Xi) and Calculate the probability distribution of that parameter
   i)Parameter may takes categorical values [Ex. parameter Soil Topography which takes value {1(Plain), 2(Gentle slope), 3(Deep slop)}]
   ii) Parameter takes continuous values, then split it depending on the probability of the range of parameter values occurred under the particular class value Si, parameter values will be partitioned into categories using ranges.[Ex. probability distribution of parameter Rainfall is as, p1(Rainfall<=38), p2(Rainfall>38)].
   Identify categorical values of parameter P

$$P=(p_1,p_2,p_3.....p_n) \quad (4)$$

3. Calculate the Entropy of P as:

$$I(P) = -(p_1*\log(p_1) + p_2*\log(p_2) + .. + p_n*\log(p_n)) \quad (5)$$

   [More the uniformity in the probability distribution, greater it gives information.]
4. T- set of records which are partitioned on the basis of class values C1, C2, .., Ck when, P- is the probability distribution of the partition (C1, C2, .., Ck) then the information needed to identify the class of an element of T is

$$Info(T) = I(P) \quad (6)$$

$$P = (|C1|/|T|, |C2|/|T|, ..., |Ck|/|T|) \quad (7)$$

5. After partitioning on the basis of class value into sets T1, T2,..., Tn
   The information needed to identify the class of an element of T is nothing but the weighted average of the information needed to identify the class of an element of Ti, i.e. the weighted average of Info(Ti):

$$Info(Xi, T) = \text{Sum for i from 1 to n of } \frac{|Ti|}{|T|} * Info(T_i) \quad (8)$$

6. Information Gain of parameter X_i is:

$$Gain(X_i ,T) = Info(T) - Info(X_i, T) \quad (9)$$

7. Repeat the procedure 2 to 6 for all the parameters to get the vector

$$G=\{Gain(X_1,T), Gain(X_2, T),....., Gain(X_n, T)\} \quad (10)$$

8. Choose the parameter X_i such that $Gain(X_i, T)$ is higher than the other parameters considered. Identify the sub-braches under that node as, If the probability distribution of the parameter X_i i.e. P is (for first iteration it will be root node) Refer Step.2
   P=(p1,p2,p3.....pn)

   Then,
   p1- subset of the dataset belongs to sub -branch 1
   p2- subset of the dataset belongs to sub -branch 2
   .
   .
   pn- subset of the dataset belongs to sub -branch n
9. Choose one of the sub-branch and respective subset of the dataset .
   Repeat the procedure 1-9 till last sub-branch will be the class node/leaf node.
10. Prune the tree if required (Logically not graphically).
11. Once the logical tree is ready verify this for the testing dataset. Error rate allowed is- Variation allowed in output: +1 level or -1 level.
    If error generated > ± 1 then, repeat the procedure 1 to 11 for the sub-branch which is identified as misclassified .

II. Model is trained using training dataset and verified using testing dataset. After model is ready the Information gain vector G in the last iteration is used for extended part b)

b) Suitability assessment matrix
Entropy/ Information gain of every parameter in all the sub branches of the resultant tree is referred from step a) equation 10. Those values are used to prepare assessment matrix. As per the ranges of information gain observed in a) and checking whether parameter is modifiable or not all $X_1, X_2, ..., X_n$ parameters are categorized into three categories. These categories ultimately shows their level of effect on output suitability level.

- High effect : rainfall, soil-type, topography and temperature.
- Moderate effect: soil moisture, EC and pH.
- Average effect: N, P, and K.

Assessment matrix is made up of five resultant vectors i.e. five rows. Lets discuss about a first row in Table 1. Let R- is a resultant vector for mapping its suitability level S1, for any specific crop (ex. jowar) for specific slot (B. Chandrasekaran, 2010), (day wise slot, one slot is of 2 months). It include suitable range for all $X_1, X_2, ..., Xn$ parameters.

$$R = \{ R_1, R_2, R_3\} \quad (11)$$

ex.
R_1 - is a set of values within a suitable range for parameters having high effect .
Under the element R_1 there are four parameters ,
X_1 - Temperature, highly suitable range R_1 = 27° to 32°C
X_2 - Soil Type, highly suitable value is Clay type
X_3 - Topography, highly suitable plane area
X_4 - Rain fall, highly suitable range is 30 to 65cm

Similarly R_2 is a vector for suitable values of the moderately affecting factors and R_3 is a vector for suitable values of the average affecting factors to achieve suitability level S1
Second row in Table 1 indicates vector to achieve suitability level S2 and so on.

Thus overall matrix for assessing suitability is dependent on categorical values as shown in Table 1.

Table 1: Suitability Assessment matrix

Suitability Levels	Levels of effecting factors		
	High	*Moderate*	*Average*
S1	H	H	H/M
S2	H/M	H/M	H/M
S3	M	M	H/M/A
N1	M	A	A
N2	A	A	H/M/A

3. Results

Analysis done by a hybrid machine learning model. Input considered is the variable parameter values from IoT based monitoring module example shown in 3.1 and three parameters from geographical information available. Output is suitability level analysed for particular crop as show in 3.2

3.1 Server side

On server side it shows the average of all input parameter values from all the nodes considered. As we know every monitoring node is going to monitor all the onsite real-time measurable input parameters trough sensors connected to it and send it to server.
Ex. Soil moisture values from four nodes as shown in fig. 4

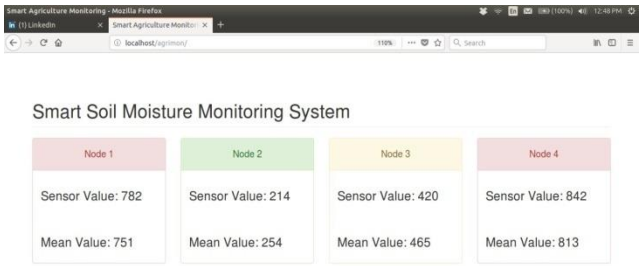


Fig. 4. Monitoring of soil moisture parameter from four nodes

Similarly all the parameters are monitored and stored at cloud. Stored data is pre-processed and used for suitability analysis

3.2 Client side

Suitability analysis is displayed in Fig. 5 for particular chosen crop using hybrid machine learning model discussed here. If lagging parameter is modifiable, then suggested to modify it by adding external components to field and improve the suitability.



Fig. 5. Suitability Analysis for crop jowar

4. Discussion

This IoT based module is capable for monitoring of all the parameters which keeps on changing. Constant parameters like soil type, topography are available in geographic data for Pune region. Historical data analysis was done in previously published work which is mentioned as a reference. It has given classification result with appropriate accuracy. So same intermediate results are used to design this assessment matrix approach and get the appropriate output with less efforts as compare to C4.5 approach.

5. Conclusion

This IoT module is a novel prototype module which can monitor multiple parameters simultaneously. Best part of this is it can made available to farmers whenever and wherever they want. Simply we can say that this is an approachable real time field monitoring system.
This hybrid machine learning model is trained once and designed output assessment matrix can be used multiple times for certain long duration. So efforts required to train the model every time are automatically reduced.

Acknowledgement

Authors are grateful to the Krishi Vidyapeeth and Meteorological Department at Pune for providing real-time dataset indirectly through publically available reports. Also grateful to VIIT, Pune and K L University for their indirect support in the completed and ongoing work.

References

[1] S. Wolfert, L. Ge, C. Verdouw, and M. J. Bogaardt, “Big Data in Smart Farming – A review,” *Agric. Syst.*, vol. 153, pp. 69–80, 2017.
[2] M. J. O’Grady and G. M. P. O’Hare, “Modelling the smart farm,” *Inf. Process. Agric.*, vol. 4, no. 3, pp. 179–187, 2017.
[3] “FAO_Chapter 6 - Causes of land degradation.” FAO, United Nation.
[4] Agriculture University, “Krishi Darshani 2008,” Pune, 2008.
[5] J. Liang, X. Liu, and K. Liao, “Soil Moisture Retrieval using UWB Echoes via Fuzzy Logic and Machine Learning,” *IEEE Internet Things J.*, vol. PP, no. 99, p. 1, 2017.
[6] R. Elsheikh, A. R. B. Mohamed Shariff, F. Amiri, N. B. Ahmad, S. K. Balasundram, and M. A. M. Soom, “Agriculture Land Suitability Evaluator (ALSE): A decision and planning support tool for tropical and subtropical crops,” *Comput. Electron. Agric.*, vol. 93, pp. 98–110, Apr. 2013.
[7] A. Benzekri, K. Meghriche, and L. Refoufi, “PC-Based Automation of a Multi-Mode Control for an Irrigation System,” in *2007 International Symposium on Industrial Embedded Systems*, 2007, pp. 310–315.
[8] M. R. M. Kassim, I. Mat, and A. N. Harun, “Wireless Sensor Network in precision agriculture application,” in *2014 International Conference on Computer, Information and Telecommunication Systems (CITS)*, 2014, pp. 1–5.
[9] R. G. Regalado and J. C. Dela Cruz, “Soil pH and nutrient (Nitrogen, Phosphorus and Potassium) analyzer using colorimetry,” in *2016 IEEE Region 10 Conference (TENCON)*, 2016, pp. 2387–2391.
[10] R. Bhimanpallewar, “A Machine Learning Approach to Assess Crop Specific Suitability for Small / Marginal Scale Croplands,” vol. 12, no. 23, pp. 13966–13973, 2017.
[11] FAO (Food and Agriculture Organization of the UnitedNations), “2. Water and soil requirements,” 2016. [Online]. Available: http://www.fao.org/docrep/u3160e/u3160e04.htm#TopOfPage.
[12] Food and Agriculture Organization of the United Nations (FAO), “Approaches to Land Classification.” p. 123, 1974.
[13] A. A. Soofi and A. Awan, “Classification Techniques in Machine Learning : Applications and Issues,” pp. 459–465, 2017.