An IOT Based Crop Selection And Productivity Enhancement System For Indian Farmers

Dhanya Viswanath1,2, Vidya Ubalanka1,3 and Anu Jose1,4

1Department of Information Science and Engineering, CMR Institute of Technology, Bengaluru, India

E-mail: 2dhanyarun0705@gmail.com, 3arnavidya@gmail.com, 4anujmattam@gmail.com

Abstract. A very large section of the Indian population depends on farming to meet their ends. But due to several reasons - natural or man-made, not limited to weather, proper irrigation mechanisms and finance, many a section of this society are often unable to reach their expected sustenance. A proper guidance mechanism that helps the farmers or aspirants select the suitable choice of crop to be cultivated in an area based on the climatic and soil conditions will increase the agricultural productivity per unit of land. In this paper, we propose a predictive system that aids the farmers in crop selection based on different climatic and soil parameters available from a dataset pertaining to Indian subcontinent. An IOT system will determine the climatic (temperature, humidity etc) and soil (pH, moisture, etc) conditions in the area of choice. This is input to the predictive system which predicts the choice of crop. The developed system can be adopted to achieve a healthy crop growth that can be adapted by agricultural practices.

1. Introduction

Agriculture, being the backbone of an economy, contributes to the economic development of a country. The scene in a developing country like India is no different. Indian agriculture accounts for close to 14 per cent of the country’s economy. Hence any development imperative to the country’s growth should begin from the agricultural industry. Technological advancements have made this achievable to an extent [1]. However, this has not reached its full potential, possibly due to the higher costs and investments involved. A substantial push will therefore be required to encourage agricultural practices to move away from strictly traditional practices and to implement new technological ways for the enhancement of agricultural output and hence the agricultural sector.

Innovative trends in Indian agriculture have seen the use of a well-defined irrigation system, use of multiple cropping and pesticides, methods enriching the nutrient content and soil fertility etc which have succeeded in increasing the yield as well as helped to bring increased revenue for the farmers. Knowing the advantages of proper soil quality and moisture, irrigation and climatic factors in the growth of a crop [2], these factors cannot be ignored and must be considered in the process of digitalizing the agricultural economy, which is much needed in the longer run. Automation would serve as a first step in this direction. Automation would ensure that proper crop monitoring is done with minimum human intervention in the field. This would be possible with the assistance of IoT.

The Internet of Things (IoT) has changed the world for the better since its advent. Along with introducing innovation in various industries, it has also helped to bring down the business operation costs. Agriculture is one such industry where IoT with its multitude of sensors and microcontrollers
has made its mark known. However, the same is yet to be pronounced true for the Indian agricultural scenario.

In this paper, we propose a system for effective selection of crop for cultivation in a portion of land based on the soil parameters and the climatic dependencies and also based on the trends in crop growth obtained from a publicly available dataset.

2. Literature Survey
In the existing Indian agricultural scenario, farmers generally decide the crop based on their past experiences growing them and based on the seasonal climatic variations in India. There is very little emphasis given to the application of technology to promote agricultural yield. This is attributed to several reasons not limited to the high initial costs involved in introducing automation in agriculture, high capital costs and a lack of awareness among the farmers about the ability of technology to revolutionize modern farm management.

There have been several works in literature focussing on analysing soil data and making predictions on various attributes affecting agriculture and soil [3]. These works have made predictions using different machine learning algorithms and taking varied parameters into account [4]-[8].

3. Proposed Architecture and Design
There are several factors affecting soil productivity. These include weather conditions like temperature and humidity, soil moisture content, pH of the soil, soil nutrient content, texture of the soil, organic matter content, water holding capacity, salinity, subsoil characteristics, and so on. The scope of this paper is limited to factors such as temperature, humidity, soil moisture content, soil pH and soil electrical conductivity (EC) and the suggested methodology has been proposed based on that.

The following is the architecture and design being proposed for incorporating the intelligence factor into Indian agriculture in order to realize an IoT based crop prediction system based on soil and weather parameters at any given location of interest. The IoT based monitoring system will be responsible for data acquisition, processing, transmission and reception functionalities. The storage functionality of cloud computing will enable sharing of resources at low economical costs, whereas the analytic capabilities of machine learning algorithms assist the farmer in analysing trends in the collected data and thereby in decision making to solve real time problems. In simple words, IoT collects the data, cloud database stores it and finally big data analyses the data to predict a plausible output.

Fig 1: Proposed architecture
3.1. The IoT subsystem

In agriculture, IoT is used to monitor and analyse various parameters affecting soil productivity. This is done with the help of sensors that work based on the wireless sensor network concept. The IoT subsystem in the proposed methodology makes use of a temperature and humidity sensor to obtain live data of the climatic dependencies affecting crop growth and a soil moisture sensor, pH sensor and a soil electrical conductivity (EC) sensor to obtain live data of the soil characteristics in the land of interest. The temperature and humidity sensor measures the temperature and humidity of the air surrounding the land. This provides the data corresponding to the climatic factors affecting soil. The soil moisture sensor measures the amount of water content in a sample of soil in the land and can be used to manage the soil moisture during the different stages of plant growth. The soil pH sensor lets us know how much acidic a soil sample is and a soil EC sensor can be used to obtain the soil variability and to indirectly determine the amount of salts present in the soil. These sensors periodically record the data from the farmer's land. Since different crops require different climatic and soil conditions for growth, the collective data from all the sensors will be helpful in choosing a crop to be grown. The data collection can be carried out real-time practically with the help of vehicle-based or vehicle-mount sensors attached to tractors that are driven over the land under investigation. This records the sensor data periodically. Another possible technique for getting these data is to make use of drones or similar surveillance mechanisms for capturing the sensor data.

3.1.1. Soil moisture sensor. The soil moisture sensor can be used to measure the amount of water present in a sample of soil. It consists of two probes which act as variable resistors and allows current to pass through to the soil sample. Water is a good conductor of electricity and hence if there is more water content in the soil, it will conduct more electricity. As a result, we get to know that the moisture content is higher. Conversely, if the moisture content is less, the soil conducts less electricity. If a predefined threshold value, which can be adjusted using a potentiometer, is exceeded, it is an indication of higher moisture content, and else it indicates low soil moisture content. LEDs present on the sensor module also help in this data capture.

3.1.2. Temperature and humidity sensor. The DHT11 is a cheap and basic digital humidity and temperature sensor that measures the surrounding air and provides the required readings with the help of a capacitive humidity sensor and thermistor. A more precise and accurate version of the sensor, DHT22 is also available, the trade-off being an increased cost, which is not preferred when looking at an economic option for farmers.

3.1.3 Soil pH sensor. Nitrogen, phosphorus and potassium are the three main nutrients which are responsible for a healthy plant growth and are also the main constituents of fertilizers generally applied to crops. But the amount of nutrients present in the soil depends on the pH of the soil, and hence it is always important to measure the soil pH and monitor it for any discrepancies. Horticultural crops have been found to satisfactorily grow in a pH of 6 to 7.5, that is, in a slightly acidic to slightly alkaline soil. However, there are also crops requiring other ranges of pH. In such cases, the soil pH can be measured and adjusted by supplementing the essential nutrients.

3.1.4 Arduino Uno + ESP8266 Wi-Fi module. Generally, microcontrollers are used in embedded systems to design and develop different devices based on the requirement. The Arduino Uno is one such open-source electronics platform mainly based on AVR microcontroller Atmega328. It consists of a USB interface, 14 digital I/O pins out of which 6 can be used for PWM output, 6 analog pins and an Atmega328 microcontroller. It can be powered using an AC or DC adopter.
Fig 2: View of Arduino Uno

The Arduino is programmed using Integrated Development Environment (IDE) software. The IDE is also required to transfer the code to the controller using the USB cable which connects it to the computer.

An ESP8266 Wi-Fi module can be used to connect the Arduino Uno to the Internet. This module is a low cost, standalone wireless transceiver that provides internet connectivity for embedded applications. It uses TCP/UDP communication protocol to connect with server/client. The Arduino microcontroller communicates with the Wi-Fi module using UART (Universal Asynchronous Receiver/Transmitter) at a specified baud rate.

Appropriate code needs to be uploaded onto the Wi-Fi module in order to connect to an existing Wi-Fi network and transfer the data received from the Arduino into a cloud server. The sensor data from the Arduino needs to be transformed into the right format for inputting into the prediction system.

3.2 ML Based Prediction System

The prediction system forms a key part of the proposed work. Once the real-time weather and soil data from the land has been collected through sensors and stored in the right format, the data is now available to be analysed. The proposed system should be capable of helping the farmer in decision making with respect to the crop to be cultivated in a specific area as per the real-time collected data. For this, we may make use of several data mining and analytical techniques which are already available. These analyse data by identifying patterns in the collected sensor data based on the training model used and suggest the crop to be grown.
4. Experimental Details
Before we begin with data collection, the prediction model must be developed. This requires the need of appropriate datasets, both for training and testing of the model. Datasets may be manually prepared taking the attributes under observation or obtained from public domain. Here, we have manually prepared the dataset, selecting soil and weather parameters as the attributes (Fig 4). The outcomes for each combination of soil and weather parameters would be the type of crop that can be grown in that condition. This method of prediction will work only if we have genuine datasets. Hence, gathering the dataset is an important factor as it forms the basis of the prediction system. This data could also be collected from various sources such as meteorological stations or from irrigation plan records that are publicly available. Once the model or system has been trained with the appropriate dataset, it should now predict the choice of crop for real time attributes. But again, the geographical area from which the dataset is obtained is important here.

We have developed the prediction model using the Weka Tool. The Weka Tool is a data mining software that has many algorithms which may be used for data analysis. Once an appropriate algorithm is selected in Weka, this is then used to finalize and generate the model. The model thus generated is then saved in our system and loaded whenever required to test live sensor data from the field. Here, we have chosen a model based on our literature survey which helped to determine which algorithm delivers more accurate results and use them for our work.

The steps in the experiment may be summarized as follows:
(i) Sample dataset collection
(ii) Partitioning of dataset into training and testing data in Weka
(iii) Choosing appropriate prediction model in Weka
(iv) Training and testing of selected model in Weka
(v) Saving finalized model in Weka
(vi) Loading the finalized model and applying live sensor data (that has already been collected) to it
(vii) Obtaining prediction on choice of crop to be grown

In our work, we have made use of a dataset which includes the temperature requirement (in degree Celsius), humidity range (in %), moisture content and soil pH of three agricultural crops namely rice, wheat and maize (Fig 5). The Weka Tool also allows us to visualize each of the selected attributes (Fig 6).
For the purpose of live data collection from the field, we require the need of sensors for capturing the attribute values [9]. Then, to analyse the data thus captured, we would need to get the data onto the prediction model (step (vi) above). For this, we make use of sensors connected to an Arduino Uno which is in turn connected with a wi-fi module that can send the sensor data to a cloud server. This live data collection is simulated using Tinkercad software, where we have made use of a temperature and humidity sensor and soil moisture sensor for obtaining weather and soil dependencies (Fig 7).

After collecting the data, it can then be formatted as per the prediction model input format requirements. Once live data is collected and the finalized prediction model has been loaded in Weka, we can now make predictions on the new data.

5. Results and Analysis
Field data is collected using Tinkercad simulations of temperature and humidity sensor and soil moisture sensor. The integrated sensor data from the Arduino is kept ready in csv format for sending to the Weka based prediction model. Different sources on crops have been made use of in getting the best dataset for training the prediction model. Weka itself allows for pre-processing of data and hence only the most required attributes from the datasets may be appropriately selected for training the prediction model. The J48, IBk and Random Tree classifiers in Weka have been used to develop the model [10]. Once the classifier is chosen, the dataset is partitioned into training and test data by selecting an appropriate percentage split. We have taken this split as 60%. This implies that 60% of the dataset will be used for training the model and the remaining for testing the model. The partitioning of the dataset into training and test datasets is a step done to test the validity of our generated model. The performance results have been compared for the J48, IBk and Random Tree classifiers based on the formulas in [10] (Table 1). The performance of each chosen classifier is depicted by the confusion matrix values. The confusion matrix combines accuracy, sensitivity, specificity and precision and generates the true positive, true negative, false positive and false negative outcomes.
Further, the model can also be more precisely tuned by including more attributes or environmental and other factors and narrowing down the options so that finally we have a fine-tuned model for crop prediction.

6. Conclusion
Crop yield prediction using intelligent machine learning techniques may improve the crop planning decisions. Different algorithms have been tested to identify the most accurate model. However, the accuracy of these different algorithms also depend on the attributes present in the dataset. In general, we have compared mostly the J48, IBk and Random forest algorithms and these have shown a higher degree of accuracy and lesser error rates over other algorithms, but these are purely dependent on the datasets being used. These diverse results for big datasets will help the farmers in predicting crops that may generate a higher yield and also help us in recommending a suitable crop to be grown.

Further development of such a productivity enhancement model will play an important role in managing the growth of crops in future by helping the farmer make the right decision at the right time with the help of technology. This would help the farmer make an informed decision about the type of crop to be grown. It will help in reducing wastage by the efficient utilization of resources and also help in optimizing the fertilizer requirements to obtain maximum productivity.

As a future scope, the data collection and processing can be further automated that would also enable the farmers to work on the data from anywhere. New modules can be incorporated in this system to monitor the performance of the farm helpers or the efficiency of the equipment involved. If the provision to predict the crop yields can be included, better planning of the crop storage, distribution and marketing can also be done from the farmer’s side.

References
[1] Rakesh Kumar, M.P. Singh, Prabhat Kumar and J.P. Singh, “Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique”, 2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), 6 - 8 May 2015. pp.138-145. M. Young, The Technical Writer’s Handbook. Mill Valley, CA: University Science, 1989.
[2] S.Bhuvaneswari, T.Pramananda Perumal, B.Jagadhesan, “An analysis and impact factors on Agriculture field using Data Mining Techniques”, IJBI, ISSN: 2278-2400.
[3] Jay Gholap, Anurag Ingole, Jayesh Gohil, Shailesh Gargade, Vahida Attar, “Soil Data Analysis Using Classification Techniques and Soil Attribute Prediction”.
[4] Ganesh P., Tamilselvi K., Karthi P., “Crop Prediction by Monitoring Temperature and Rainfall Using Decision Tree with Iot and Cloud Based System”, IOSRJEN, ISSN (p): 2278-8719.
[5] R.Harine Rajashree, C.Kavya, T.Kiruthika, J.Nisha, “Crop Recommendation System for Precision Agriculture”, 2016 IEEE Eighth International Conference on Advanced Computing.
[6] Y. Xiaoxia, Z.Chengming, “A soil moisture prediction algorithm base on improved BP”, 5th International Conference on Agrogeoinformatics (Agro-geoinformatics), 2016.
[7] J. Dong, S.C. Steele-Dunne, T. E. Ochsner, N.van de Giesen,”Determining soil moisture by
assimilating soil temperature measurements using the Ensemble Kalman Filter”, Advances in Water Resources, 2015, 86, 340-353.

[8] A. Kumar & N. Kannathasan, (2011), “A Survey on Data Mining and Pattern Recognition Techniques for Soil Data Mining “, IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 3

[9] Dr M Suchithra, Asuwini T, Charumathi M C, Ritu N Lal, “Sensor Data Validation”, IJPAM, ISSN: 1314-3395.

[10] Kusum Lata, Khan, Sajidullah S, “Experimental Analysis of Machine Learning Algorithms Based on Agricultural Dataset for Improving Crop Yield Prediction”, IJEAT, ISSN: 2249 – 8958.