Crop Recommendation System

Pradeepa Bandara
Lecturer
Sri Lanka Institute of Information
Technology
Sri Lanka

Thilini Weerasooriya
Assistant Lecturer
Sri Lanka Institute of Information
Technology
Sri Lanka

Ruchirawya T.H.
Undergraduate
Sri Lanka Institute of Information
Technology
Sri Lanka

W.J.M. Nanayakkara
Undergraduate
Sri Lanka Institute of Information
Technology
Sri Lanka

Dimantha M.A.C
Undergraduate
Sri Lanka Institute of Information
Technology
Sri Lanka

Pabasara M.G.P
Undergraduate
Sri Lanka Institute of Information
Technology
Sri Lanka

ABSTRACT

Automating agricultural aspects is a mechanical process with or without human intervention in agriculture. Due to less space of domestic lands, it has become an important area of choosing the most suitable crops based on prevailing factors in the selected area. In Sri Lankan even though there are enough knowledge, techniques, and methods which are done manually available in agriculture, there is not any system in which the environmental factors are detected and suggests the user which crop type is best for farming. This paper is consisting of a theoretical and conceptual platform of Recommendation system through integrated models of collecting environmental factors using Arduino microcontrollers, Machine learning techniques such as Naïve Bayes (Multinomial) and Support Vector Machine (SVM), Unsupervised machine learning algorithm such as K-Means Clustering and also Natural Language Processing (Sentiment Analysis) concerned with the Artificial Intelligence to recommend a crop for the selected land with site-specific parameters with high accuracy and efficiency. It has been a major problem to identify what to grow, any man has adequate space in the owner's land. Not only domestic lands but also for farming lands. Why it has become a problem is that environmental factors such as temperature, water levels, and soil conditions are uncertain as they change from time to time. Due to these problems, this solution of crop recommendation system predicts the user, what crop type would be the most suitable for the selected area by collecting the environmental factors for plant growth and processing them with the trained sub-models of the main of the system.

General Terms

IOT – Internet of Things

Keywords

Agriculture, Machine Learning, Arduino, Natural Language processing, Farming

1. INTRODUCTION

Modern farms and agricultural operations ar taken place more totally different than those many decades agone, primarily due to advancements in technology, as well as sensors, devices, machines, and knowledge technology. Today's agriculture habitually uses refined technologies like robots, temperature and wetness sensors, aerial pictures, and GPS technology, and lots of complex IOT devices. These advanced devices in agriculture enable businesses and farmers to be additional profitable, efficient, safer, and more environmentally friendly.

The rise of digital agriculture and its connected technologies has opened a wealth of latest knowledge opportunities. Remote sensors, cameras, and alternative connected devices will gather data twenty-four hours per day over a complete farm or land. These will monitor plant health, soil condition, temperature, humidity, etc. the quantity of information these sensors will generate is overwhelming. this enables farmers to achieve a far better an improved understanding of state of matters on the bottom through advanced technology which will inform them additional regarding their situation more accurately and quickly.

The environmental data that is gathered by remote sensors are processed by algorithms and statistical data which will be understood and helpful to farmers for decision makings and keep track of their farms. The more inputs and statistical data collected, and higher the algorithmic rule is at predicting the outcomes. And the aim is that farmers will use these technologies to attain their goal of improved harvest by creating better selections within the field. By implementing the system of temperature, soil hydrogen ion concentration and soil wetness detection, the information captured are processed with an explicit algorithmic rule and passed to a centralized database that is connected to different modules of the research, so the main system will predict the most effective crop kind that the farmer should grow to require the most outcome of the crop kind that is farmed in a home garden or the respectable land area.

2. LITERATURE SURVEY

In this literature review of this project, the team sought out and studied various patents, research papers, documents, and newspapers and magazine articles from various scenes. The paper [1] states requirements and why they tend to move into precision agriculture [2] which is due to globalization are discussed. Precision agriculture is site-specific farming. Though Precision agriculture has shown an improvement with time, there exist some issues. As mentioned above site-specific methods of such systems are needed to be supervised to get an improved result. Only a few of the outcomes are provided a particular result. Nevertheless, the situation is farming is indispensable since if any default or a mistake occurs, it might lead to serious damage to resources and as well as the plants.

In this research, it is proposing a system where the major factors are taken into consideration at the same time and come up with a solution so that the system will not be complicated for the user. As mentioned above in the sentence, the major factors taken into consideration at once is unlike other models proposed in previous researchers, this system considers all the major factors which are essential for plant growth, are processed together using various algorithms whereas the other models consider only parameters at once keeping the other factors constant.

As an example, some tests are carried out to find the rate evaporation and how the plant growth is affected when there's no sufficient water present. And a derived equation is presented as a result [3]

ETo = K pan × E pan

ETo: reference crop evapotranspiration

Kpan: pan coefficient E pan: pan evaporation

Even though an equation is proposed, there are some limitations [4]. Mainly this could be done only for a land with a less area. This is not suitable for commercialization as the profit will be low when you use less area for cultivation. The second limitation is the average rainfall of Sri Lanka mostly suitable for many crop types to grow without any deficiency. Therefore, the water level alone itself is not a big issue whereas other factors are. If there's enough water for plants but no temperature that would be a problem as the major environmental factors have a mutual relationship with each other in plant growth.

The ultimate goal of the previous researches is also to predict the best crop type. But once the farmer or user has cultivated the predicted crop type, the system's job is done. But in the system that proposed from this paper, has a feedback system as well. Even after suggesting the best crop type, the system can track the plant growth and it provides feedback if the farm is malnourished. So that the user can take necessary precautions prior.

3. METHODOLOGY

3.1 Dataset Collection

It is needed to have a certain temperature, humidity, soil pH, sunlight, and soil moisture for a plant to be grown healthy. To receive to good harvest those conditions should be satisfied. But those conditions may vary according to the plant varieties. The initial data set is collected from the Department of Agriculture Sri Lanka [5], other agriculture books, Agricultural web sites [6], and other reports and research papers. This initial data set was used to train the crop recommendation model so that the accuracy would be increased.

Table 1. Sample data set

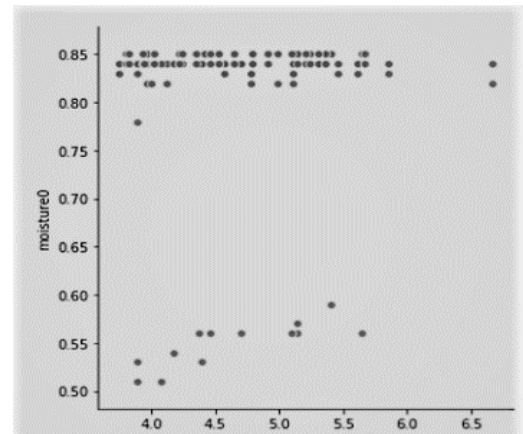| Crop	Temp.	Hum.	pH	Sun.	Moist.
Bean	26	63	7	8	63
Lettuce	23.8	60	6.5	8	60
Carrot	26	98	6.5	8	98
Cabbage	29.4	95	7.1	5	95
Beet	30	99	6.5	6.5	6.5

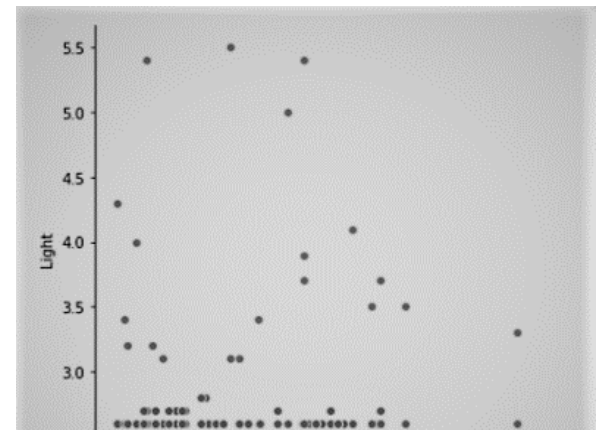**Fig 1: Data plot of soil moisture using Linear Algorithm**

Fig 2: Data plot of Sunlight (lux) using Linear Algorithm

3.2 Collecting Environment Factors

To compare and predict the initial data set, the environmental factors needed to be gathered. To collect environmental factors, Arduino microcontrollers [7] are used. Since both the temperature and humidity sensors are comprised of a single micro-controller, four sensors are used to collect data. They are sunlight intensity sensor, soil moisture sensor, soil pH sensor, and humidity and temperature sensor. Those sensors are connected with an Arduino Wi-Fi module and gathered data is sent to the database. The gathered data are cleaned and processed by using clustering and other algorithms to pass the values to the next component of crop recommending and stored in the database.
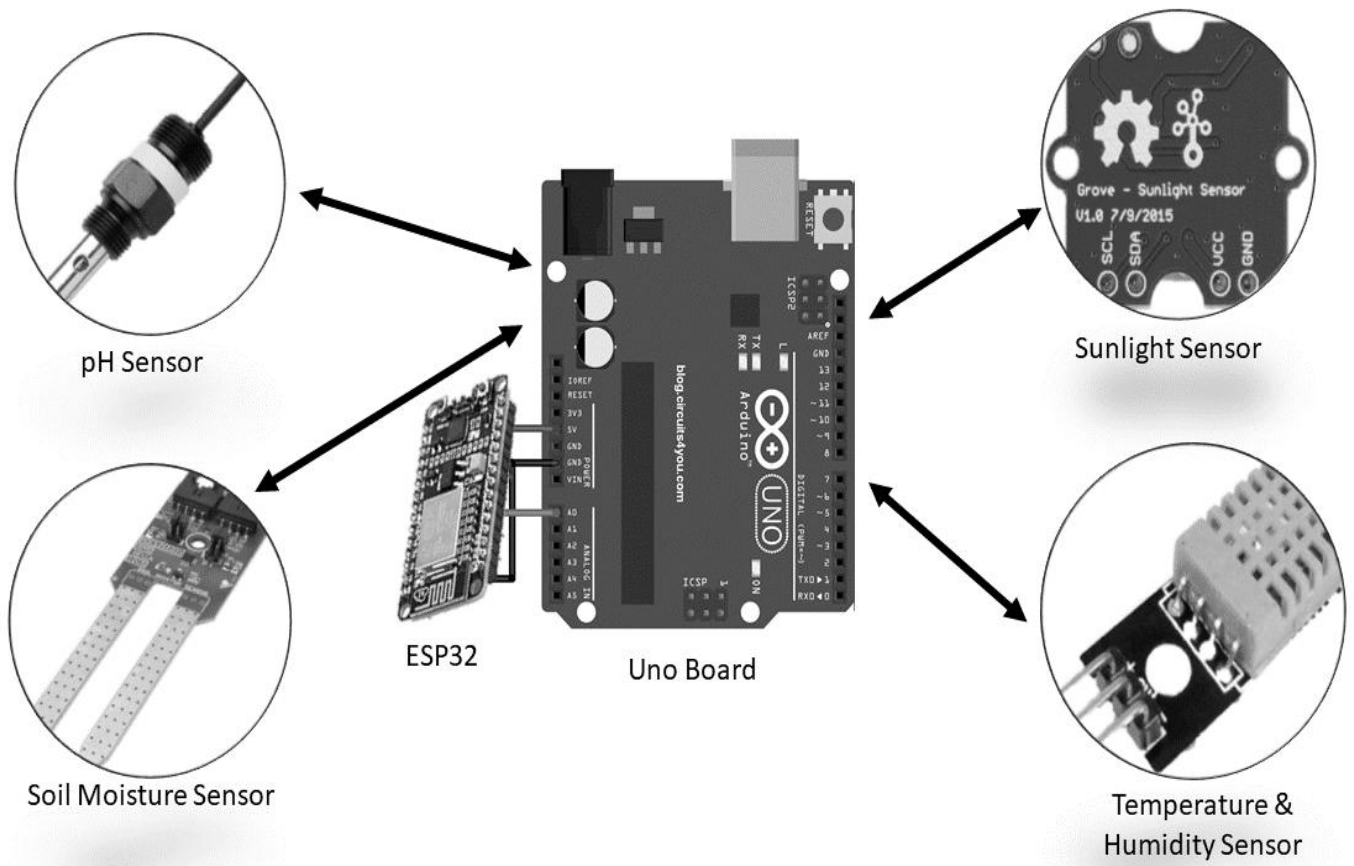
Fig 3: Arduino Micro-Controllers Connectivity

3.3 Crop Prediction

Since the environmental conditions differ from region to region, a machine learning model is used to predict the best crop type for the selected land. To train the crop recommending model with the data collected from the Arduino sensors, machine learning algorithms [8] are used to identify the best crop to cultivate with the highest probability of growing. Naïve Bayes & Support vector machine algorithms are used to select the best crop type.

From this model, it decided what type of crops that the farmer should grow. This is done by analyzing factors of humidity, temperature, soil moisture, pH level, and sunlight. Mainly the system suggests 4 crop types by analyzing the above-mentioned factors using two machine learning algorithms.

Naïve Bayes [9] - Naive Bayes is a technique for constructing classifier models that assign class labels to problem instances which are represented as vectors of feature values, where the class labels are drawn from some finite set.

Support Vector Machine (SVM) [10] - The objective of the support vector machine algorithm is to find a hyperplane in N-dimensional space (N — the number of features) that distinctly classify the data points.

3.4 Monitoring and Feedback

This proposed system product would mainly identify four types of crops according to the environmental factors of the selected plot of land. But the reason to obtain a probability of more than 90% for the above-identified crops would be the soil condition or any other changes in the selected land. But to avoid these factors affecting the crop prediction, the farmer's feedback system is included in the system.

Once the crop type is recommended, the farmer is asked for details and feedbacks regularly through the mobile application to guide the farmer with necessary precautions. To provide the necessary feedback, the feedback system is used in the mobile application by selecting the crop type. By this, the overall accuracy and the reliability of the product are increased with time.

4. RESULT

The proposed system was successfully implemented in a selected land. The data sampled from the sensors at an interval of one hour on a normal sunny day is presented in Table I. Based on the tests conducted and the data collected, it can suggest the best crop to cultivate for maximum harvest. Using the feedbacks collected from the farmer the accuracy of the predictions is sharpened by neglecting the invalid data. As an example, if the farmers are continuously providing the negative feedback for cultivating strawberry in Galle, the system itself learned, and the accuracy of the final output is increased which means in future it will not suggest strawberry to grow in Galle. The farmer can enter his own feedback using his native language. Since the most Sri Lankans use Sinhala and English, the custom libraries are implemented, by which the system identifies the language significantly by using natural language processing algorithms and the feedback is identified.

According to the prevailing environmental factors in the selected land, the best suitable four crop types are suggested to the farmer.
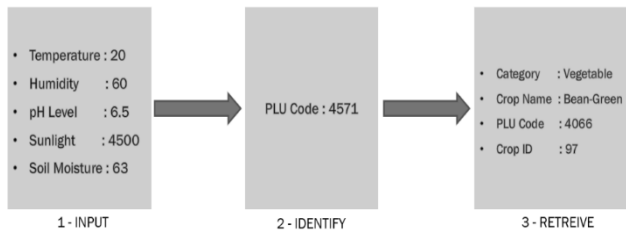

Fig 4: A sample output given to the farmer

Figure 4 shows a sample result where the farmer gets, once the environmental factors are entered. PLU code is an attribute that is used to identify the crop uniquely.

The overall accuracy of the proposed system is more than 92%. As the farmer continue to use this, the more data system feeds, the more accurate it will be. Therefore. by the time of long usage, the farmer can obtain more the 95% of accuracy from the whole system.

Table 2. Accuracy level of the models in the system

Name of the Module	Accuracy
Temperature and Humidity detection model	92%
2Sunlight intensity, soil pH and soil humidity detection model	95%
Crop recommending system	90%
Monitoring and feedback system.	96%

5. CONCLUSION

In a modern environment with less space and less knowledge of agriculture, all the factors are considered from the perspective of farmer and plant, and the farmer is properly guided until the harvesting. Before selecting any plant to grow it is important to have the knowledge and an understanding of the factors that affect the cultivation and how to maintain or control them. From this system, these above-mentioned factors are automatically processed and select the crop type to be cultivated.

Once the plant is cultivated, the farmer is asked for feedback regularly with a time interval of one month. From this feedback taken, the system self-trained, and the accuracy is improved with time and data collected. From this system, the guidance of a specialist is not needed, and the maintenance is less. Thus, implementing this system will not have any additional monetary impact on the user.

Since the average land space for a person in Sri Lanka is less, from the data gathered by sensors it is evident that this proposed system has above 95% accuracy and suitable for both rural and urban areas. With the time of the accuracy would be more increased significantly as the significant amount of data is collected and the system self-trained by the time of approximately one year. The experiments conducted over the 4-6 months proved the reliability and accuracy of the system. When the whole world is looking forward for new agriculture consultancy measures, this system will definitely be a torchbearer.

6. FUTURE WORKS

It can do a lot of additional features to the system. Currently, it takes necessary environmental factors as inputs and suggests a very suitable crop to be cultivated. But as the next level, the Automation part can be added as the response system to the feedback. This can be modified to control the humidity, water level, etc. according to the need of the farmer.

Presently it takes all environmental factors as inputs to the system, but as an additional feature, an algorithm can be implemented to predict the one factor using another two factors. (Example – predicting the soil pH level from soil moisture and sunlight), so that the initial cost of setting up the sensors would be less and can be easily maintained.

7. REFERENCES

[1] Lakshmi.N, Priya.M, Sahana Shetty, and Manjunath C. R, Crop Recommendation System for Precision Agriculture, vol. 6 Reading, IND: International Journal for Research in Applied Science & Engineering Technology, 2018. [Online] Available: www.ijraset.com.Ding, W. and Marchionini, G. 1997 A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.

[2] Remi Schmaltz, "What is Precision Agriculture", April 2017. [Online]. Available: https://agfundernews.com/what-is-precision-agriculture.html [Accessed Feb.23, 2020].

[3] C. Brouwer and M. Heibloem, Irrigation Water Management: Irrigation Water Needs, manual 6 Reading, ITALY: Food and Agriculture Organization of the United Nations, 1987. [Online] Available: http://www.fao.org/3/s2022e/s2022e07.htm#TopOfPage.

[4] A. de Carbon, "PRECISION AGRICULTURE: ITS BENEFITS AND LIMITATIONS", carrhure, 2019 [Online]. Available: https://www.carrhure.com/precision-agriculture-benefits-limitations/ [Accessed: Feb.25, 2020].

[5] Department of Agriculture Sri Lanka (2015), 'Crop Suitability Recommendation for Grama Niladhari Divisions in Sri Lanka', Natural Resources Management Center, Peradeniya.

[6] Marie ionnotti (2020), 'Outdoors & Gardning', The Spruce Web Site

[7] WolkWriter (2019), 'Periodically measure environmental conditions and send them to WolkAbout IoT Platform to monitor the environment remotely.', ARDUINO PROJECT HUB

[8] Simon Tavasoli (202), '10 Machine Learning Algorithms You Need to Know', Simplilearn Solutions

[9] Sunil Ray (2015), '6 Easy Steps to Learn Naive Bayes Algorithm with codes in Python and R', Analytics Vidhya

[10] Rohith Gandhi (2018), 'Support Vector Machine — Introduction to Machine Learning Algorithms', Toward Data Science Web Site