CROP YIELD PREDICTION USING IoT AND MACHINE LEARNING

Jit Saha[1], Arka Bhowmik[1], Srija Talukdar[1], Shakyajyoti Pyne[1], Shayari Mondal[1], Mrs. Moumita Goswami[2]

[1]Student, CSE Department, MCKV Institute of Engineering, Liluah, West Bengal

[2]Assistant Professor, CSE Department, MCKV Institute of Engineering, Liluah, West Bengal

jit291999@gmail.com arkabhowmik007@gmail.com pyneshakyajyoti007@gmail.com s5050mondal@gmail.com srijatalukdar7@gmail.com moumita02314@gmail.com

Abstract. Agriculture is the primary factor that is important for the survival and the economy in India. It is important to increase the productivity of agricultural and farming processes to improve crops and cost-effectiveness. Crop Yield Prediction using IoT and Machine Learning is a management concept focused on providing the agricultural industry with the infrastructure to leverage advanced technology including big data, cloud, and the internet of things (IoT) – for tracking, monitoring, automating and examining the operations. Farming can be made more productive and accurate with the help of technologies like IoT. IoT can be used in different fields of farming or agriculture. The implementation area of IoT is wide and can be implemented in every field. The Internet of things (IoT) is a promising technology that provides efficient and reliable solutions for the modernization of several domains. IoT-based solutions are being developed to automatically maintain and monitor agricultural farms with minimal human involvement. Furthermore, IoT-based agriculture systems' connection with relevant technologies including cloud computing, machine learning, and analytics makes it more vibrant. In this paper, we have developed the system, that will help us to collect the relevant data from the different agricultural fields with the help of the sensors (i.e. Dht-11 sensor, Soil-moisture-water sensor, Rainfall sensor). Then this collected data is uploaded to a cloud server and the uploaded data is analysed with the dataset available by using the Random Forest machine learning algorithm and then the suitable result is displayed on the user's device.

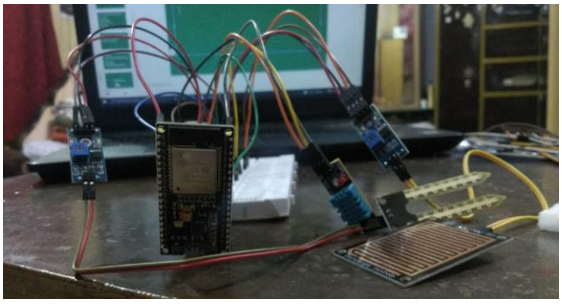
Keywords: Smart Agriculture, IoT, Sensors, Machine Learning, Crop prediction.

1. INTRODUCTION

To improve the agricultural yield with fewer resources and labor efforts, large inventions have been made throughout human history. Cultivating maximum crops at minimum cost is one of the goals of agricultural production. Early detection of the management of problems associated with crop yield indicators can helpful in increasing the yield and subsequent profit [1]. However, the high population rate never let the demand and supply match during all these times. According to the forecasted figures, in 2050, the world population is expected to touch 9.8 billion, an increase of approximately 25% from the current figure. Almost the entire mentioned rise of the population is forecasted to occur among the developing countries. The trend of urbanization is also projected to continue at an accelerated pace, with about 70% of the world's population predicted to be urban by 2050 (currently 49%). Moreover, income levels will be multiples of what they are now, which will drive the food demand further, especially in developing countries. As a result, these nations will be more cautious about their diet and food quality; hence, consumer liking can move from wheat and grains to legumes and, later, to meat. In order to feed this higher, more urban, and richer population, food production should double by 2050. Particularly, the current figure of 2.1 billion tons of annual cereal production should touch approximately 3 billion tons. Not only for food, but crop production is becoming equally critical for industry; indeed crops like cotton, rubber, and gum are playing important roles in the economy of many nations. Additionally, the food-crops-based bioenergy market started to increase recently. Even before a decade, only the production of ethanol utilized 110 million tons of coarse grains (approximately 10% of the world's production). Due to the rising utilization of food crops for biofuel production, bio-energy, and other industrial usage, food security is at stake. These demands are resulting in a further increase of the pressure on already scarce agricultural resources, resulting in immediate improvement in the agricultural yield. In this paper, we have developed an Arduino device that will detect the necessary climatic conditions like - soil moisture, rainfall, humidity and temperature using sensors like DHT11 and Rainfall sensor, and will upload the readings data to the Firebase cloud and then use the machine learning model Random Forest algorithm, to detect the crop which is best suited for the particular land and display that predicted crop to the farmer in his mobile phone using Bluetooth.

2. LITERARY SURVEY

This literature survey has been evolved around the different discussions related to various patents, research papers, documents and magazine articles from various scenes [2]. Crop recommendation system predicts and tells 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 system. Site specific management is used. Though Crop Recommendation has shown an improvement with time, there exist some issues. As above mentioned site-specific methods of such systems are required to be supervised to get an improved result. Only



a few of the outcomes provide a particular result. Nevertheless, the situation is that farming is indispensable since if any default or a mistake occurs, it might lead to serious damage to resources and as well as plants. The Internet of Things (IoT), the design of getting real-world objects connected with each other, will change the ways we establish, obtain and consume information radically. Agriculture can be connected to the IoT through the sensor networks and which allows us to create connections among agronomists, farmers and crops regardless of their geographical variances. With the help of the connections, the agronomists will have a better understanding of crop growth models and farming practices will be improved as well. This paper describes on the design of the sensor network when connecting agriculture to the IoT. Different parameters are considered in the design like Reliability, management, interoperability, low cost and commercialization. Finally, we are sharing our experiences in both development and deployment of the system [3]. 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, the major factors taken into consideration at once is unlike other models proposed in previous researches, this system considers all the major factors which are essential for plant growth, and are processed together using various algorithms whereas the other models consider only parameters at once keeping the other factors constant. Keeping in mind the farmer's financial condition, and the availability of network connection in the area, the perfect crop is recommended to the farmer via bluetooth, otherwise crop recommendation could also have been done via wifi as ESP32S Node MCU supports both wifi and bluetooth. Internet of Things (IoT) is a new paradigm where almost each and every physical object can be furnished with ability of sensing, communicating, and processing capabilities allowing them to communicate with other devices possibly via the internet and/or any sort of telecommunication networks.[4]

PROPOSED SYSTEM

This device reads the climatic as well as the soil conditions of the farm based upon the different kinds of sensors like temperature, humidity, soil moisture, and ph. It provides the recommended crop to the farmer based on the prediction done by the Classification algorithm.

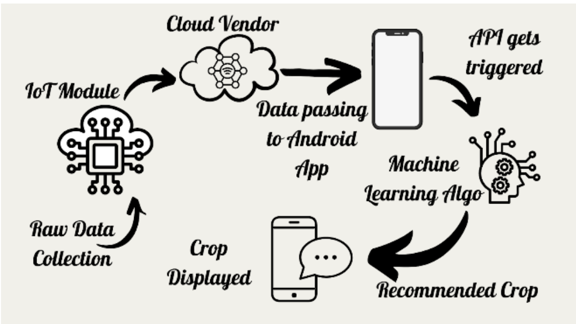


Fig 1. Workflow Diagram of The Proposed System

In the above workflow diagram, the raw data of the climatic and soil parameters like soil moisture, pH, humidity, temperature and rainfall are collected from the fields by the IoT device using the sensors. The data is then uploaded to the firebase cloud as shown in the workflow. Then the data is passed onto the android application, which has a ML model at the backend. The backend is a flask-based python application which acts like an API, taking the data from the app as a request and recommending a suitable crop as a response. Python pickling is being used to save the Random Forest model behaviour and python unpickling is used to load the pickle file whenever it is required[5].The prediction takes place based on the random forest algorithm, which gets displayed to the user end, over the android application.

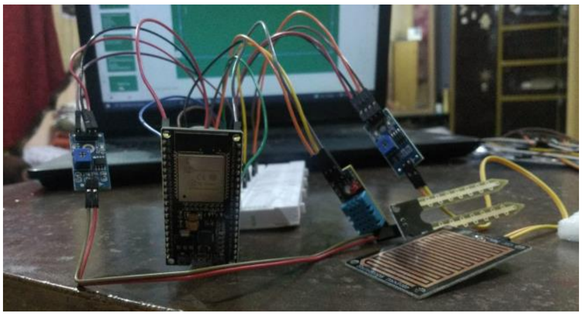


Fig 2. Proposed System

The above proposed IoT model (Fig 2), in association with the machine learning algorithm, determines the appropriate crop based on the soil and climatic conditions. Random Forest algorithm is used as it is better than linear regression when the data has high SNR (Signal to Noise Ratio).

3.1. Data Collection

The vital part of a supervised learning project is the training data. In agriculture, predictive analytics the dataset helps to predict or identify the soil nutrients level required for the crops like Rice, Wheat, Jute etc.,[6].The data acquired for training a model is very important to predict a resultant output as per the pattern predicted from the provided input. The data is collected from various databases which shows the case study of the last five years of data and re-authenticated with our Arduino module. The size of our dataset is 7000 x 5(i.e.7000 rows and 5 columns). The rows consist of 70 different crops under 100 different conditions each and the columns of our dataset are temperature, humidity, rainfall, pH and Label.

```
array(['Rice', 'Maize', 'Banana', 'Jute', 'Pulses', 'Mango', 'Papaya',
       'Tea', 'Aleovera', 'Arecanut', 'Ashwagandha', 'Rose', 'Blackgram',
       'Chickpea', 'Coconut', 'Coffee', 'Cotton', 'Grapes', 'Kidneybeans',
       'Mothbeans', 'Mungbeans', 'Muskmelon', 'Orange', 'Piegonpeas',
       'Pomegranate', 'Watermelon', 'Apple', 'Cabbagge', 'Cauliflower',
       'Green Chillies', 'Carrot', 'Ginger', 'Garlic', 'Onion', 'Brinjal',
       'Button Mushrooms', 'Potato', 'Capsicum', 'Tomato',
       "Lady's Finger", 'DragonFruit', 'Olive', 'Marigold', 'Beetroot',
       'Lettuce', 'Corn', 'Green Peas', 'Cucumber', 'Guava', 'Turmeric',
       'Rajma', 'Pumpkin', 'Litchi', 'Broccoli', 'Spinach', 'Groundnut',
       'Jackfruit', 'Radish', 'Chinese Cabbage', 'Drumstick', 'Soybean',
```

Fig 3. Crops of the dataset

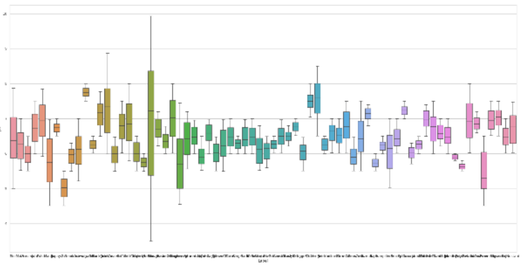


Fig 4. Box plot of the dataset

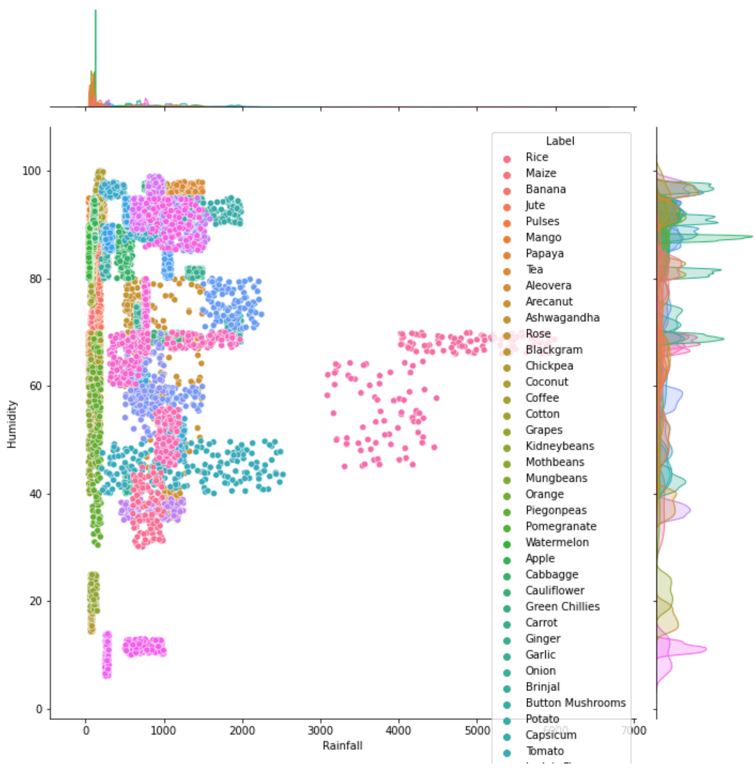


Fig 5. Join plot of the dataset

3.2. Algorithm Used

Random Forest is one of the most popular machine learning algorithms which belongs to the supervised learning techniques. It can be used both for the Classification problems and Regression problems in machine learning. It is based on the concept of the

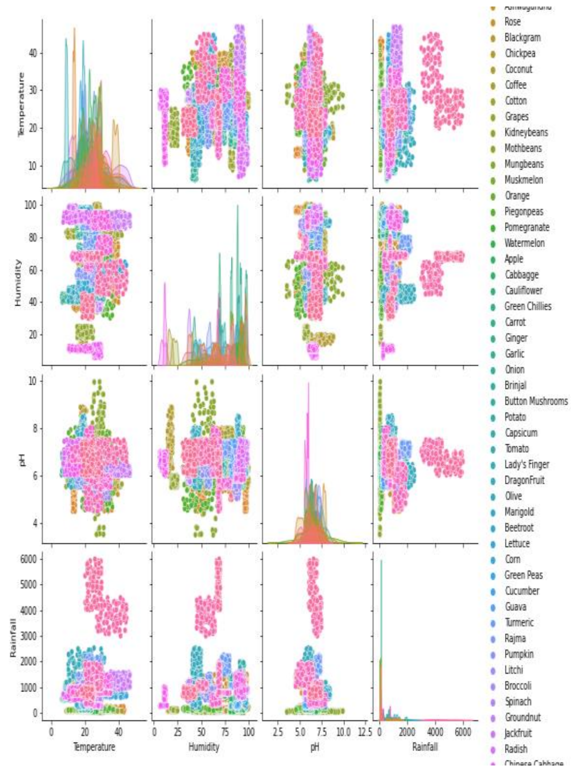


Fig 6. Pair plot of the dataset

ensemble learning which is a process of combining multiple classifiers to solve a complex problem and also to improve the performance of the model. In this paper, the crop prediction problem is solved by proposing a recommendation system through an ensemble model with majority voting technique using Random tree.[7]

The higher number of trees in the forest leads to greater accuracy and prevents the problem of overfitting in the model.

3.2.1 Assumptions for Random Forest

As the Random Forest combines multiple trees to predict the class of the dataset, it is highly possible that some decision trees will predict the correct output, whereas others will not. But putting together, all the decision trees it will predict the correct output. Hence two assumptions have to be made for making a better Random Forest classifier:

- Some actual values must be there in the feature variable of the dataset so that the classifier can predict accurate results rather than a guessing the result.
- There must be very low correlations from the predictions of each tree.

3.2.2 Why use Random Forest?

Some points are given below to explain why we should use the Random Forest algorithm:

- The training time is less compared to other ML algorithms.
- The output prediction is correct and is very high even in the case of large datasets
- It maintains the accuracy even when a large proportion of data is missing.
- Random Forest algorithm has an ability to analyse the growth of crop related to the current climatic conditions and biosphere.[8]

3.2.3 How does Random Forest algorithm work?

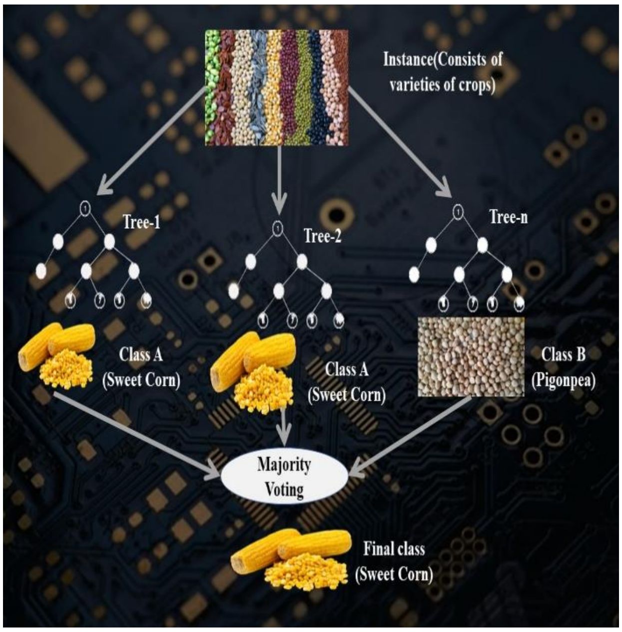


Fig 7. Working of Random Forest on our dataset

In our project the collected raw data from the field using the IoT device which has been sent to the Firebase is now used by the Random Forest model to predict the crop. The data collected gets compared with the instance which is the database containing the crops under different conditions. Now n numbers of decision trees are generated with their respective outputs of recommended crop. By using majority voting method the crop which gets predicted by the most number of decision tree is considered as the final predicted crop and is displayed as the output.

3. AUTHOR'S CONTRIBUTION

Author 1: Jit Saha

IoT model design and cross platform connection

Author 2: Arka Bhowmik

Machine learning model deployment

Author 3: Shakyajyoti Pyne

Android App Backend development

Author 4: Shayari Mondal

Android App Front End development

Author 5: Srija Talukdar

Flask API deployment and Firebase Integration

4. RESULTS

Using the Random-Forest the accuracy of our Crop Recommendation is 96.5%. The Android application developed by us displays the crop to be grown on the field based on the field parameters collected by our IoT device. On top of that sensor nodes are also capable of sending this data to cloud. An Android application has been developed in order to access all these agricultural parameter.[9]

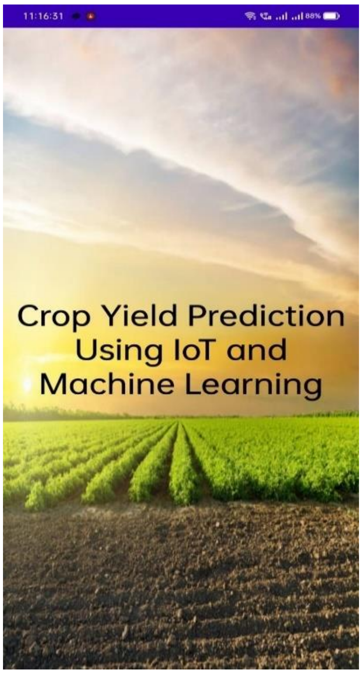


Fig 8. Home screen of the Android App

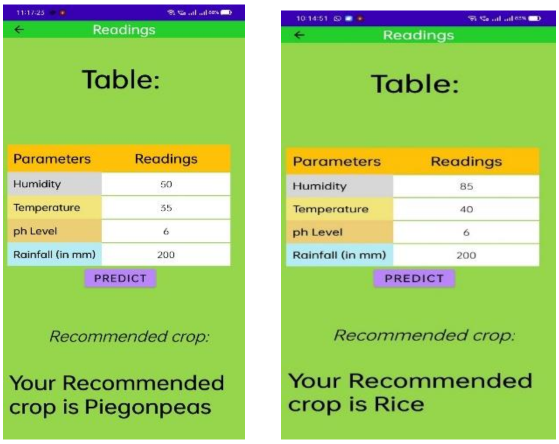


Fig 9. Results of recommended crops as per the data collected

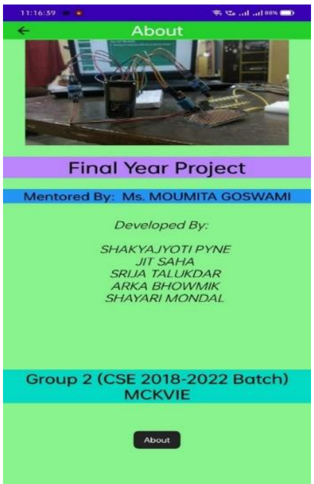


Fig 10. Developer's name screen of the app

Fig 11. Real time data collection firebase

6. CONCLUSION AND FUTURE SCOPE

India is a nation in which agriculture plays a prime role. In the prosperity of the farmers, the nation prospers. This work would help farmers in sowing the right seed based on soil and climate requirements to increase productivity and acquire profit from such a technique. Thus, the farmer can plant the right crop increasing his yield and also increasing the overall productivity of the nation. Our future work is aimed to have an improved data set with a large number of parameters and also implement the crop prediction[10].

Overall, the project has achieved its objectives, our IoT device is successfully working and the Random-Forest algorithm is giving appropriate results accordingly. Thus the device can be used to predict the best fit crop to be grown in the fields.

The future scopes of our project include-

a) Fertilizer recommendation with the quantity measures.

b) Acknowledge the farmers about various loans and policies which will be helping them to earn extra profit.

c) Implementation of various cost-effective technological measures like push message notification using Bluetooth technology over our current module which would lead to no mandatory need for smart phones.

d) Implementing other possible machine learning algorithms to check the models efficiency and increase the accuracy if possible.

7. REFERENCES

[1] Miss.Snehal S.Dahikar (EXTC), Dr.Sandeep V.Rode (EXTC), "Agricultural Crop Yield Prediction Using Artificial Neural Network Approach" International Journal Of Innovative Research In Electrical, Electronics, Instrumentation and Control Engineering (Vol.2, Issue 1, January 2014)

[2] Pradeepa Bandara, Thilini Weerasooriya, Ruchirawya T.H., Dimantha M.A.C, Pabasara M.G.P, "Crop Recommendation System", International Journal of Computer Applications (0975 – 8887) Volume 175– No. 22, October 2020.

[3] Junyan Ma, Xingshe Zhou and Zhigang Li, "Connecting Agriculture to the Internet of Things through Sensor Networks", Internet of Things (iThings/CPSCom) 2011 International Conference on and 4th International Conference on Cyber Physical and Social Computing, pp. 184-187, Oct 2011.

[4] Boulos Wadih Khoueiry; Mohammad Reza Soleymani "A Novel Machine-to-Machine Communication Strategy Using Rateless Coding for the Internet of Things" IEEE Internet of Things Journal (Volume: 3, Issue: 6, Dec. 2016)

[5] Akkem Yaganteeswarudu "Multi Disease Prediction Model by using Machine Learning and Flask API" 2020 5th International Conference on Communication and Electronics Systems (ICCES).

[6] M. Chandraprabha, Rajesh Kumar Dhanaraj "Soil Based Prediction for Crop Yield using Predictive Analytics" 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)

[7] S. Pudumalar; E. Ramanujam; R. Harine Rajashree; C. Kavya; T. Kiruthika; J. Nisha "Crop recommendation system for precision agriculture".2016 Eighth International Conference on Advanced Computing (ICoAC).

[8] V. Geetha; A. Punitha; M. Abarna; M. Akshaya; S. Illakiya; A.P. Janani "An Effective Crop Prediction Using Random Forest Algorithm" 2020 International Conference on System, Computation, Automation and Networking (ICSCAN)

[9] Manishkumar Dholu; K.A. Ghodinde "Internet of Things (IoT) for Precision Agriculture Application" 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)

[10] S.Pudumalar, E.Ramanujam, R.Harine Rajashreeń, C.Kavyań, T.Kiruthikań, J.Nishań. , "Crop Recommendation System for Precision Agriculture", 2016 IEEE Eighth International Conference on Advanced Computing (ICoAC)