IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⅀ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11417

Rice Crop Yield Prediction using Machine Learning and Integrating IoT

Praveen Kumar Arjun Patel[1], Ojas Ajit Daware[2], Deep Dulal Debnath[3], Prathmesh Shirish Haware[4]

[1-4]University of Mumbai BVCOE, Navi Mumbai Maharashtra, Pincode- 410210

Abstract: This paper focuses on predicting the rice harvest and investigating the factors affecting rice production in various regions of the Maharashtra region of India. The software aims to provide a rice harvest using a random forest algorithm and accurately predicting the yield. To demonstrate the effectiveness of harvest forecasting, an Indian government database will be used in 34 districts of the Maharashtra region, India [2]. Boundaries such as rainfall, temperature, humidity, and location are given as a contribution to the random forest model to define the annual variation of the regional rice crop in Maharashtra. The software will also use other IoT devices to retrieve real-time data from the field. This will give an accurate result to farmers and prevent major losses to farming. With the help of powerful services like Amazon Web Services(AWS), Java, and Flask it tends to work on low-end devices and remote regions.

Keywords: Amazon Web Services(AWS), Flask, REST API, Android Studio, MySQL, Java, Arduino Ide and Random Forest Regression.

1. INTRODUCTION

Predicting crop yields is critical to global food production. Policymakers rely on accurate predictions to make decisions during import and export to improve national food establishments. Considering that yield is determined by the interaction between management, soil, and climate, the purpose of the harvest should vary from season to season, but also from region to region. Many crop yielding models use mathematical or simulation models. This paper proposes a software system that uses machine learning techniques and concepts that will provide an effective way to predict crop yields under a variety of crop conditions. Using additional technologies such as IoT, using sensors in all fields to monitor real-time results, and integrating with Machine Learning will provide better accuracy. Making them on a large scale and including other vegetables and rice can be beneficial to the whole country.

2. OBJECTIVES

1. Provide accurate forecasts for rice yields using Machine Learning.
2. Integrating real-time data obtained from various sensors.
3. Using data obtained from climate and soil sensors to predict conditions and timers for farmers. Existing System.

3. EXISTING SYSTEM

3.1 Existing systems.

1. Existing systems now use ANN structures to predict paddy yields when considering one weather parameter [1]. Some studies were conducted in large corporations on several farms of the same company.

1. In this study, both satellite data and field-based surveys were used to map crop acreage and yield forecasts using SNAP software.

2. The data sets used in the previous study were taken from public records of the Government of India from 1998 to 2002.

3. One proposed plan was to match the retrospective model that would help us predict crop yields as a function of monthly temperature and rainfall. The system can use regression because we need to predict the yield based on historical yield data compared to the environmental parameters.

4. In the system comparison model, the MLR is a method used to analyze the causal and impact relationships between the objective (y ≡ Yield) and the predictive variables used.

5. In the system, KNN (k-neighbor) was used to store all available cases and classify new cases based on the similarity scale. KNN has used both separation and retrospective problems [4].

6. In some systems, the Deep Gaussian system for predicting the data output of remote sensor data was used.

Some systems have used the Spiking Neural Networks method for Crop Yield Estimation Based on Spatiotemporal Analysis of Image Time Series.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁜ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11417

3.2 Comparison of the existing system

3.2.1 Rice Crop Yield Prediction

As discussed, previously different systems used different machine learning algorithms to predict the crop yield of rice crops. Using climatic conditions as the limits of crop yield crops were predicted in various existing systems. For example, applying the weather and land conditions simultaneously will not only give the rice crop condition but will also analyze the soil conditions necessary for a good yield and provide the necessary solutions to improve the crop yield of rice in case the yield is lower than average.

3.2.2 Allowing Farmers to use the site for result

Considering the previous system, these were designed as a research project and did not allow farmers or users to fully participate in the program. It

easy for the user to get the guessing results the way he or she likes. After the user has obtained the result, he can determine if his yield is low and find the solution needed to improve the condition of his farm and other factors that may be able to increase their productivity and improve their yield.

Soil Moisture contains water in the soil. Capacitive Soil Moisture Sensor is the best and most affordable sensor to measure soil moisture. The Earth's analog sensor output can be converted to percentages and can be displayed at 0.96OLED.

4. PROPOSED SYSTEM

This paper works on developing a system that will accurately predict the yield of a rice crop by providing additional information on the necessary soil conditions. This will help you save time and give better results.

4.1 Implementation Details

4.1.1 Random Forest Regression

Random forests or random decision forests are an integrated learning method for classification, retrieval, and other activities, which works by creating several decision trees during training and class removal which is a classroom (planning) or direct prediction (retrieval) of individual trees. Random decision-making forests correct the practice of deciduous trees to complete their training set. In this study, we selected the number of decision trees to be 100 and the amount of data to be taken as a sample of 80% of the data. The Bagging meta-algorithm is used to make every tree that is produced independently of pre-built trees. This is achieved by pressing samples. Once the model is trained, we use the remaining 20% of the database to test our model.

4.2 Analysis

The algorithms used are the Random Forest Regression. To demonstrate the effectiveness of harvest forecasting, an Indian government database will be used in 34 districts of the Maharashtra region, India. Boundaries such as rainfall, temperature, humidity, and location are given as a contribution to the random deforestation model to define the annual variation of the regional rice crop in Maharashtra. R squared values are usually between 0 and 1 (which can be negative if the model is worse than the horizontal line). A value close to 1 indicates that the return line is very similar to the data. Also, we have set a data set in the Reverse plan to emphasize patterns in the database during the analysis of the test data. Reversal passages as the name suggest form a retrospective line between 2 parameters and help visualize their linear relationships.

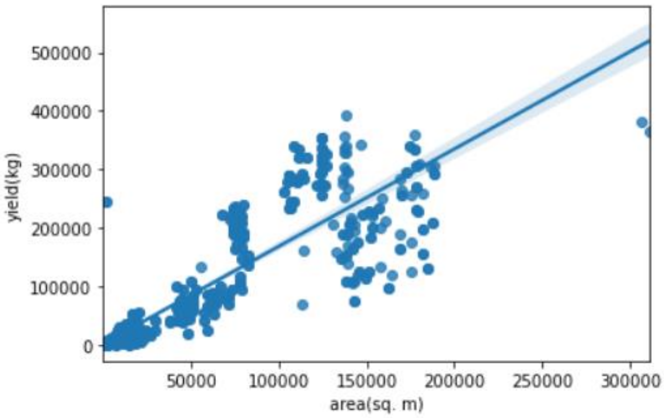

Fig 1: Reg Plot

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⁙ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11417

4.3 IoT Integration

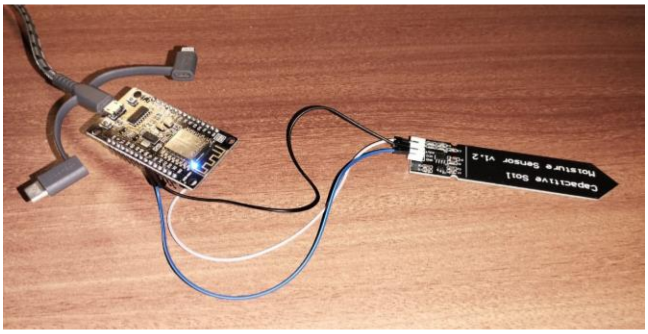

Fig 3: IoT Sensors

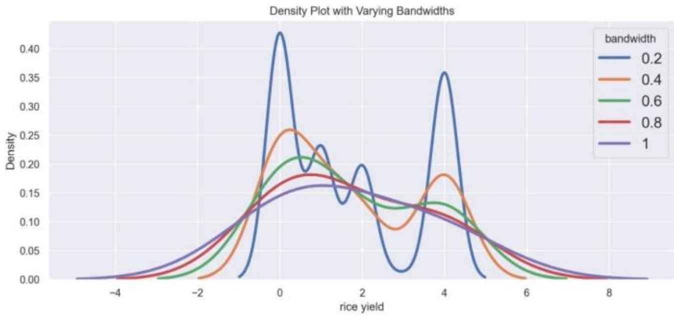

Fig 5: Density Plot

Fig. 5 shows the density plot which represents the distribution of numeric parameters that are humidity and yield. It visualizes the distribution of data over a continuous interval of time.

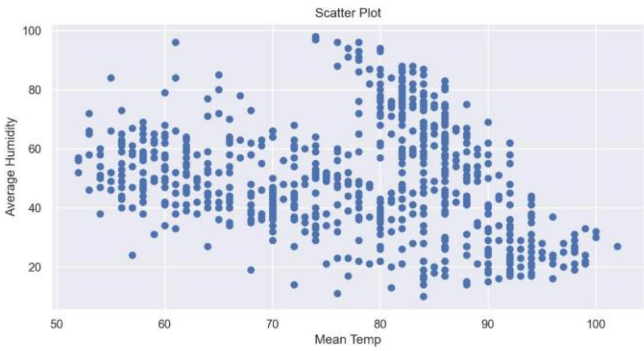

Fig 6: Scatter Plot

Fig. 6 plot diagram uses Cartesian coordinates to display values for typically two variables for a set of data. It shows whether the two parameters have any correlation or not.

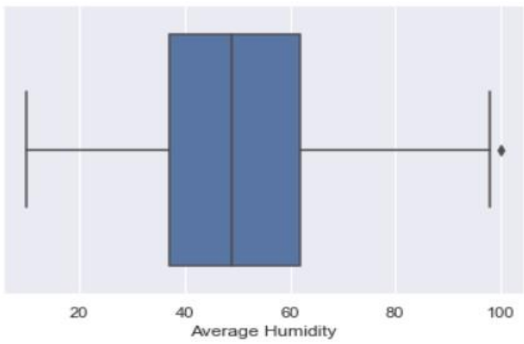

Fig 7: Box Plot

Fig. 7 shows a standardized way of displaying the distribution of data based on the parameter.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 7.39 ⁘ Vol. 11, Issue 4, April 2022
DOI: 10.17148/IJARCCE.2022.11417

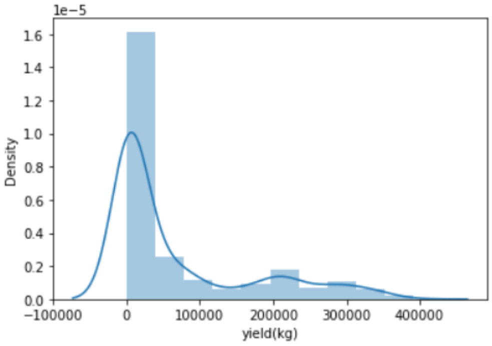

Fig 8: Density Plot

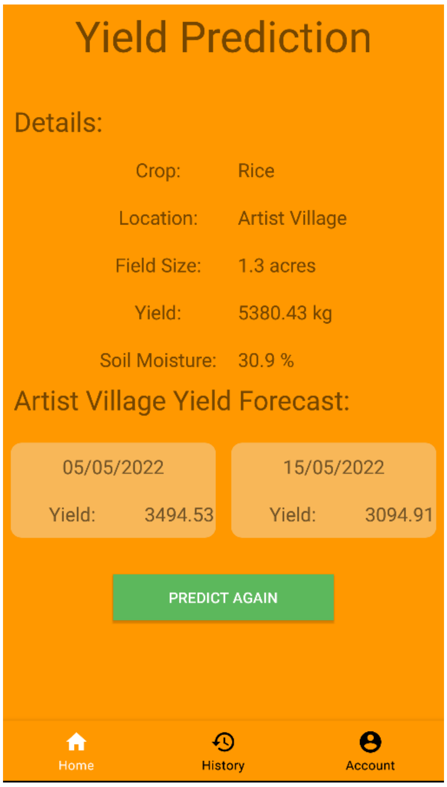

Fig 10: Rice Crop Yield Prediction

After entering all the details like crop name and field size, the application shows the JSON result of the rice crop yield prediction predicted using Random Forest Algorithm and IoT sensors. Along with the current prediction, it shows 10 days and 20 days forecasted results of the yield.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11417

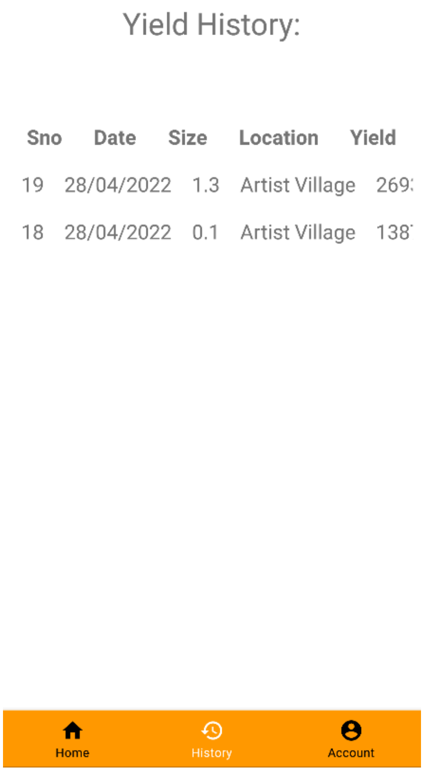

Fig 11: Yield History

Along with predicting results, it also stores the results in the Amazon Relational Database service in MySQL format where we can easily retrieve them and display them in the application.

4.4 IoT Devices

4.4.1 Soil Moisture Sensor

This sensor measures the amount of water in the soil and gives it the right amount of moisture as it comes out. This soil moisture sensor measures soil moisture levels by capacitive sensing rather than the opposite sensor like the sensors on the market. It is made of corrosion-resistant materials which give it excellent service life. Apply it to the soil around your plants and monitor real-time soil moisture data. This module includes an onboard voltage controller that provides a working range of 3.3 ~ 5.5V. Ideal for a low-voltage microcontroller with both 3.3V and 5V power supplies.

4.4.2 NodeMCU ESP8266

NodeMCU is a Lua-based firmware-based open source and development board specifically targeted at IoT-based Applications. Includes firmware running on ESP8266 Wi-Fi SoC from Express if Systems, as well as module-based ESP-12 module.

4.4.3 Arduino IDE

Open-source Arduino Software IDE makes it easy to write codes and upload them to the board. This software can be used with any Arduino board. The IDE application is suitable for different operating systems such as Windows, Mac OS X, and Linux. It supports the programming languages C and C++.

4.4 Connection Diagram

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 7.39 ⋇ Vol. 11, Issue 4, April 2022
DOI: 10.17148/IJARCCE.2022.11417

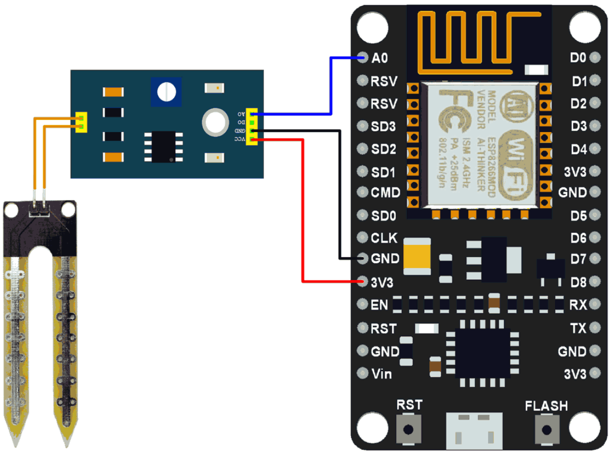

Fig 4: Connection Diagram

5. RESULTS

All the objectives and the implementations which are mentioned in the proposed system are implemented. The screenshots are displayed in the current section. The screenshots are taken from the android application, cloud server running in Amazon Web Service, and MySQL workbench.
The application is built using Java, Flask, Arduino, and Amazon Web Service. It is a microservice application where the services are loosely coupled. Various AWS services used are Amazon Simple Storage Service, AWS Lambda, Amazon Elastic Compute Cloud, Amazon Rekognition, Amazon Relation Database Service, Amazon DynamoDB, AWS IoT Core, AWS Amplify, etc.

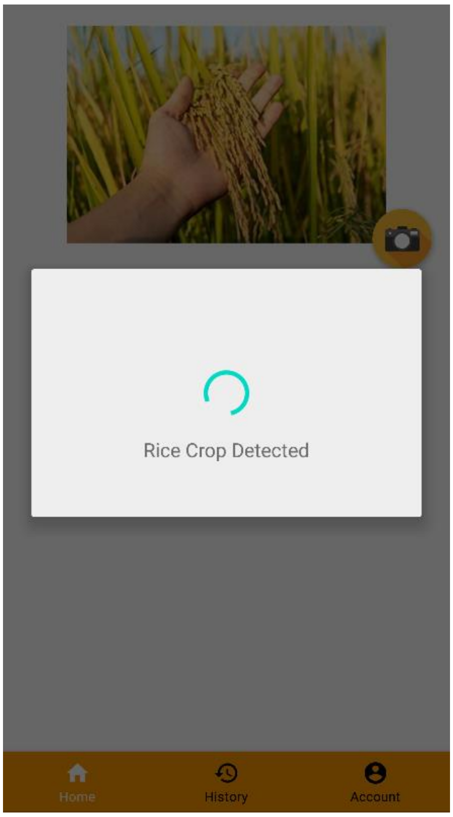

Fig 8: Image Detection

Before predicting, first, we need to capture the image of the crop, or we can choose the image from the gallery. Then the prediction is made on the image using the Amazon Rekognition service. According to that, the crop is detected.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 7.39 ⋇ Vol. 11, Issue 4, April 2022
DOI: 10.17148/IJARCCE.2022.11417


```
{
    "currentyield": {
        "date": "24/04/2022",
        "location": "Artist Village",
        "size": 0.1,
        "yield": "1645.53",
        "soil moisture": 32.5
    },
    "forecast10": {
        "date": "04/05/2022",
        "yield": "2062.20"
    },
    "forecast20": {
        "date": "14/05/2022",
        "yield": "1582.34"
    }
}
```


Fig 9: JSON Response to API Request

The application makes a REST API call to the AWS Elastic Compute Cloud instance where the rice crop yield prediction code is running and the incoming request is served by Flask. After the successful prediction, it sends JSON data back to the application which then displays the results.

ricecrop §

```
#include "FS.h"
#include <ESP8266WiFi.h>
#include <PubSubClient.h>
#include <NTPClient.h>
#include <WiFiUdp.h>

// Update these with values suitable for your network.

const char* ssid = "[redacted]";
const char* password = "[redacted]";

WiFiUDP ntpUDP;
NTPClient timeClient(ntpUDP, "pool.ntp.org");

const char* AWS_endpoint = "a[redacted].amazonaws.com"; //MQTT broker ip

void callback(char* topic, byte* payload, unsigned int length) {
Serial.print("Message arrived [");
Serial.print(topic);
Serial.print("] ");
for (int i = 0; i < length; i++) {
Serial.print((char)payload[i]);
}
Serial.println();
```

Fig 12: Arduino IDE

The Arduino IDE is used to connect NodeMCU ESP8266 with AWS IoT Core using MQTT Protocol. For creating a thing in IoT Core certificates and policies are required. Sketch is used to upload the certificates to the device's drive rather them storing them inline in the script.

▼ outTopic April 24, 2022, 00:18:52 (UTC+0530)


```
{
  "Id": "11",
  "soil moisture": "30"
}
```


▼ outTopic April 24, 2022, 00:18:50 (UTC+0530)


```
{
  "Id": "10",
  "soil moisture": "30"
}
```


Fig 11: Data uploaded to IoT Core

The data that is sent through the NodeMCU are uploaded to IoT Core where it is stored in NoSQL Database.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 7.39 ⅔ Vol. 11, Issue 4, April 2022

DOI: 10.17148/IJARCCE.2022.11417

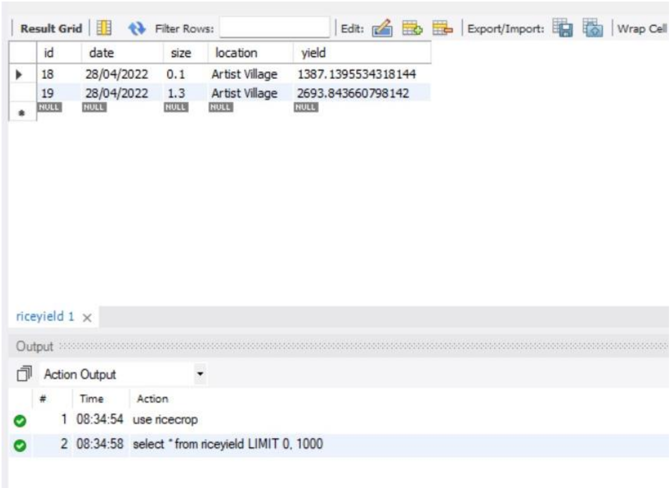

Fig 12: MySQL Workbench Result

6. CONCLUSION

In recent years, great strides have been made in the challenging task of predicting the harvest of rice. Developing accurate models for crop yield estimation using different technology may help farmers and other stakeholders improve decision-making about national food import/exports and food security. This research has demonstrated the prediction of rice crop yield by applying machine learning techniques. Hence, it concludes that classifiers used on the current study dataset and reported earlier should be recommended for further development of a rice prediction model.

7. REFERENCES

[1] Vinushi Amaratunga, Lasini Wickramasinghe, Anushka Perera, Jeevani Jayasinghe, and Upaka Rathnayake. 2020. Artificial Neural Network to Estimate the Paddy Yield Prediction Using Climatic Data.
[2] Niketa Gandhi, Owaiz Petkar, Leisa J. Armstrong. 2016. Rice Crop Yield Prediction Using Artificial Neural Networks.
[3] Ayush Shah, Akash Dubey, Vishesh Hemnani, Divye Gala and D. R. Kalbande. 2018. Smart Farming System: Crop Yield Prediction Using Regression Techniques.
[4] Girish L, Gangadhar S, Bharath T R, Balaji K S, Abhishek K T. 2018. Crop Yield and Rainfall Prediction in Tumakuru District using Machine Learning.
[5] Jiaxuan You, Xiaocheng Li, Melvin Low, David Lobell, Stefano Ermon. 2017. Deep Gaussian process for crop yield prediction based on remote sensing data.
[6] Pritam Bose, Nikola K. Kasabov, Lorenzo Bruzzone, Reggio N. Hartono. 2016. Spiking Neural Networks for Crop Yield Estimation Based on Spatiotemporal Analysis of Image Time Series.