IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

IJARSCT

Impact Factor: 6.252

Volume 2, Issue 5, May 2022

Farmer's Friend: An AI Model to Predict Crop Yield and Provide Suggestion

Thombare Shubhangi Radhakisan[1], Zaware Payal Eknath[2], Dongare Snehal Anil[3], Kasar Yogesh Shivaji[4]

UG Student, Department of Information Technology[1,2,3,4]

Amrutvahini Polytechnic, Sangamner, Maharashtra, India

Abstract: *A framework for agricultural yield forecasting and fertilizer suggestion is amongst the most critical demands of the moment. Water shortages as well as other soil degradation have grown increasingly prevalent in recent years, causing enormous loss of human life and livelihood all across the world. Crop output unpredictability has increased as a result of increasing incidences of global climate change and fluctuating weather. Due to the unpredictability, agricultural productivity has likewise been difficult to anticipate. This necessitates the development of a crop yield forecast and fertilizer suggestion method, which has been accomplished in this research article with minimal error. As a result, in order to improve the process of agricultural production prediction, this study examines the deployment of machine learning technologies. Linear clustering, Artificial Neural Networks, and Decision Making are all used in this research's technique. The technique has been assessed for the existence of any errors that have resulted in good operational outcomes.*

Keywords: Linear Clustering, Artificial Neural Network, and Decision Making, etc.

I. Introduction

Agricultural techniques have always been among the most important factors in a country's economic success. Food is grown in accordance with the country's need, allowing the people to be supported successfully at a minimal cost. Indigenous agriculture is a practical and successful strategy that enables for the implementation of better practices. Agriculture is critical to every country's development. It has a significant impact on the country's economic progress. Utilization of adequate soil is a crucial requirement for increasing agriculture productivity.

In this scenario, soil testing is a critical component of improving farming efficiency. Soil testing is now done manually, which takes a very long time and introduces the danger of human error. As a consequence, soil testing has become inevitable in today's world. This research uses a novel technique to soil testing that is simpler to implement, takes relatively little time, because it is less costly. The prognosis for soil diversity has a significant influence on agriculture's potential to improve output. To boost productivity, agriculture production must be strengthened. As a response, producers must do a soil analysis prior to developing a harvest. The pH of the soil influences the amount of fertilizer applied in a given plantation area.

Biologically, medically, and chemically, soil micronutrients and pH are important factors in overall agricultural productivity. Crop-specific requirements such as soil phosphorus, potassium, and nitrogen can be satisfied by adding the deficient element. The soil characteristics within a field might change based on the season, previous crop, and other variables. As a consequence, a soil quality inspection is advantageous for producing healthy plants. Soil is the earth's crust's outer covering, which envelopes the planet's surface and is used for farming.

The primary purpose of soil testing and evaluation of samples taken from a given area is to provide a comprehensive understanding of soil features, productivity, and bioactive components, as well as suggestions for enhancing the soil's overall quality. In regards to effectively implementing the substantiation obtained from the results and evaluations of soil measurement metrics, it is essential that the connection between soil nutrients and their control on the advancement of various plant variants be well comprehended, while taking other factors in the area into account.

Soil analysis also helps to improve long-term soil nutrition and overall soil conditions by supporting sustainable practices and other natural resource conservation in order to keep the soil fertile while minimizing the detrimental effects of synthetic substances. Soil is a complex mixture of geological, biochemical, and organic components. The organic parts are made layers of dead and living fauna and vegetation, while the elemental arrangement is the result of weather induced breakup of the substrate incorporating structural and thermal causes. The agricultural sector makes a considerable contribution to India's GDP. Economic conditions, catastrophic weather occurrences, management of human resources, and labour expenses are all possible stumbling blocks to the agricultural sector's growth. Farming productivity, on the other hand, must be raised as the world's population grows.

To reduce manual soil management, a variety of sensors are being utilized to assess a variety of soil properties, such as moisture content, surface temperature, and availability of nutrients, such as nitrogen, phosphorus, and potassium. According on the observations of the soil properties, the kind of soil is predicted. Farmers can use the suggested strategy to avoid utilizing the current soil testing procedure rather than using the appropriate fertilizers for their crops. Farmers are yearning for new agricultural methods and technology that will allow them to enhance total output.

In section 2 of this research paper, relevant studies are discussed. In section 3, the suggested approach is described in detail. In part 4, the experimental assessment is carried out, and section 5 closes this study piece with the potential for future improvement.

II. Literature Review

Supachai Puengsungwan [1] According to the report, ecologically sound electricity is essential for human survival in the foreseeable. This article discusses how solar panel solutions and Internet of Things (IoT) technologies may be used to lower solar panel infrastructure costs for hydroponic intelligent agriculture. The core application of the internet of things-based leaf sensor is the evaporation idea, which is used to achieve the study goal. One of the most noticeable advantages of the Internet of Things is the ability to communicate in timely manner between crops and growers. In particular, for optimization, the Wi-Fi based controller may regulate the hydraulic motors with in network. The energy consumed of hydroponic lettuce cultivation using the suggested leaf sensing control system is greatly lowered, per the test result.

Dhiraj Sunehra [2] explains how numerous sensors are used to construct a smart urban agricultural system. The Raspberry Pi is used to monitor and transmit the sunlight intensity, moisture levels, dampness, and heat to the website. As a result, this system demonstrates the significance of IoT networks. These services are growing increasingly important because they allow equipment to be controlled through the internet. The technique that has been devised effectively delivers a supply of food to all of the community surrounding an area, making urban farming more productive and economical.

Byunghyun Ban [3] outlines a completely autonomous fertilization system's architecture and operating algorithm. This approach is intended for use in controlled hydroponic systems. Instead from monitoring pH and EC, the method measures specific ion levels. This system's sole aim is to provide insufficient ions with highest precision. It contains an entirely automated ISE calibration mechanism, as well as various nutrition concentrates for calibration. ISE artefact control techniques are completely interoperable with the system. This solution makes use of an Azure-based cloud environment to circumvent the high costs of on-premise solutions that are currently available. The mechanical component and automated system functioned well, monitoring and regulating chemical constituents of water and nutrient containers even without presence of a plant.

Tairan Xiong [4] suggests that the approach achieved its aim of minimizing agricultural variety monitoring while also lowering the price and resource requirements of traditional human interventions. At a fraction of the expense of modern computing hardware, it has equivalent operating and preventive capacities. Even during its early stages of development, the architecture solved a bevy of existing issues and developed a slew of innovative accessories. For the first time, sophisticated biometric identification algorithms are being employed in agricultural farming. The system's adaptability and architectural adaptability make it an appealing implementation choice.

Nahina Islam [5] describes the characteristics of LPWAN solutions and assesses their applicability for precision agriculture in isolated regions. When comparing different LPWAN techniques, it is clear that LoRaWAN is particularly applied for an IoT process that requires low energy consumption, high capacity, and extremely maneuverable end devices with a wide scope of information. As a result, the transmission range of LoRaWAN controllers and the energy consumption model of LoRaWAN end devices have been given. Researchers offer a specific example of an intelligent sprinkler that employs an IoT system backed by the LoRaWAN communication protocol to fully appreciate the capability of LPWAN for distant precision agriculture.

Shiva R [6] Propose that comprehensive robotized agriculture on a small scale can produce food products. Androids can be used in three different ways to complete the task. For starters, plants are produced in batches with consistent multiple rows that are easy to track. The three robots used are one for putting chemicals, collecting, the second for evaluating plant state, and the third for sowing, all of which are simply drones operating in a customized method. The robotic arm is attached to a moving autopilot system. The microprocessor, transceiver, camera, ultrasonic sensors, and power source are all included in this automobile. When the broadcaster sends the information, the vehicle's microcontroller receives it via the receiver and adjusts the speed of the motors to move against the batch. The transmitters are placed in precise spots around the field and are in charge of managing the robotic vehicle's operation.

Nelson Bore [7] AGWS is a blockchain-based electronic field records (EFR) platform that contains data acquired through multiple technologies like as longitudinal information gathering, mobile-based data collecting apps, IoT equipment, and weather remote monitoring, among others, to codify and electronically represent fields. AGWS was developed in close partnership with the small-scale agricultural community, which included intensive seminars and field trips, as well as interviews with different ecosystem stakeholders. A huge number of additional farm activities events were analysed, and machine learning techniques were constructed to provide reliable electronic field records (EFR) and autonomous farm boundaries recognition. To meet the aims of the suggested solution, all of these are integrated with the AGWS blockchain network.

Heechang Chung [8] explains how precision agriculture is one of the most pressing demands of the hour in terms of incorporating innovation and technology into the farming system. The majority of farmers are not technologically savvy and need to be taught how to use certain aspects of precision agriculture. Smart farming education program distributes farming information to farmers and reports farmers' problems to the related educational system for resolution to assist farmers in making farm management decisions. This is a smart farming training program that helps farmers gain expertise in the proper management of fields.

Harshkumar Prakashbhai Thakor [9] states that the farmers in the Indian subcontinent are following farming practices that are very old and traditional in nature. These techniques are tried and tested over the years and are guaranteed to work, but these approaches are time consuming and leads to a lot of redundancy for the farmer. Therefore, this approach needs to be improved further to facilitate the development of highly efficient agriculture that can save a lot of effort and expenses for the average farmer. Therefore, the authors in this approach devise a model for the analysis and the deployment of a smart digital farming approach that optimizes the agricultural production.

Xiaofan Jiang [10] indicates the importance of agricultural productivity and growth for the economic growth of the nation. Farming is a key activity in which day-to-day crop surveillance is required in order to maximize crop development and production. With the increase in consumption for food, the area of farmland that has to be supervised by a producer has increased significantly. As a result, contemporary techniques and equipment are required to attain enhanced yield type and effectiveness. The researchers have designed a very efficient wireless sensor network made up of flexible nitrate monitors. These sensors give vital data via the LoRa framework, allowing for significant increases in the amount and reliability of the yield.

III. Proposed Methodology

The presented technique for the purpose of achieving crop management and a yield prediction through the use of Artificial Intelligence and Internet of Things platform has been illustrated in the system overview shown in figure 1 above. The proposed approach has been attained through the implementation of a number of steps that are described below.

Step 1: Dataset, Sensor Data Collection and Pre-processing

A Desktop standalone application is developed for performing the analysis of the farm. The user will take the developed application to destination farm along with the PH sensor and Arduino Uno Microcontroller for interfacing the same. The user will contact the farmer and manually add the NPK or nitrogen Phosphorus and Potassium values of the soil and utilizes the PH sensor by inserting it into the field. The PH sensor will record the PH values and send it to the Arduino Uno microcontroller to record these values to be utilized for evaluation in the subsequent steps by the java code.

The presented technique utilizes input data collected from the farm using sensors along with a corresponding dataset extracted from the URL: https://github.com/VaibhavSaini19/Crop-Yield-Predictionusing-ML. This input dataset is beneficial in process of yield prediction that is going to be performed in this methodology.

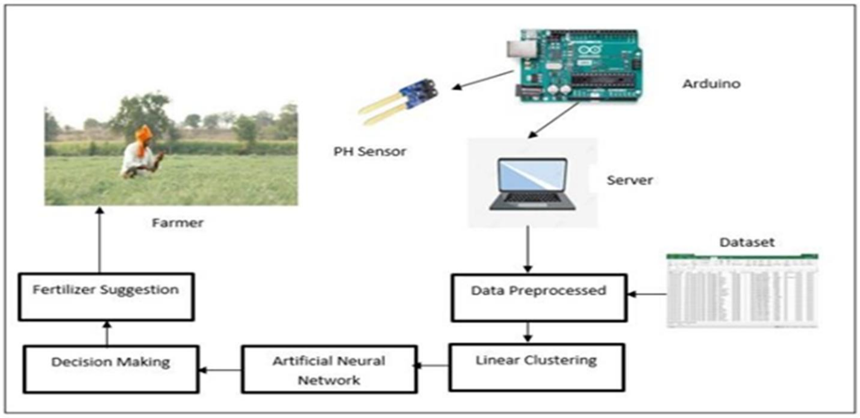

Figure 1: Proposed Model System Overview

The retrieved dataset is fed into the presented approach, which is used to make accurate yield predictions. The dataset is tightly integrated in the form of a worksheet, which the system reads using the JXL API. The capability to interact with the workbook file is provided by this package to Java code.

The dataset has a vast amalgamation of attributes, but just a few are taken into account for our approach because they can help us achieve a much precise yield forecast. The field area, yield, temperature, precipitation, humidity, soil type, crop type, and district are all retrieved parameters. These are the appropriate attributes; the remaining attributes are redundant and will be removed. These characteristics are vital in defining the forecast and have proven to be successful in achieving an increase in production, which may be quite advantageous.

The user needs to register into the system before gaining access to the features of the standalone application. The user utilizes the Graphical User Interface for the purpose of creating the login credentials by providing the system various attributes such as User ID, password, mobile number, email address, gender, Date of birth and name. When the relevant details are provided to the system by the user for successful registration, these values are stored in the database. The user can now gain access to the system by providing the respective login credentials. The authenticated user provides the respective parameters for the NPK values as well as the readings for the PH values of the soil. These values are provided to the next step for the.

Step 2: Linear Clustering

In this step of the procedure the previous step provides the input in the form of a double dimension list. This list consists of the selected attributes that are pre-processed and labelled. The clustering on these attributes is achieved through the use of the linear approach.

The user attributes that have been entered through the user interface is also taken as input in this step of the procedure. This input is collaborated in the form of a list which is there subjected for distance evaluation from the attributes and the rows of the data set. For the purpose of distance evaluation distance has been utilized through the equation 1 given below.

$$ED = \sqrt{(x1 - x2)\,2 + (y1 - y2)\,2} \text{-------} (1)$$

Where, x1, x2, y1, y2 are the entities of the Rows.

ED - Euclidean distance of a specific row

The attributes of the data set which are preprocessed previously are also taken as an input in the form of a list and each one of these values are subjected to distance calculation with the user attributes. After the calculation of the distance each subsequent row is updated with the distance called the row distance at the end of each row. This procedure is performed for all the rows in the data set and the resultant list is effectively sorted in ascending order of these row distances.

The top 50% of the sorted road distance list is selected as the inner list and the bottom 50% is considered as the outer list which results in the clustering of the data set effectively with respect to the user input provided to the standalone application. The achieved clusters are segregated in two distinct list which are provided to the next step for the purpose of implementation of the artificial neural networks.

Step 3: Artificial Neural Network

This is the most crucial and important step of the proposed approach. This type utilizes the data achieved in the previous steps for the purpose of estimating the crop yield and the deficiency in the nutrients of the soil. In this step of the procedure the labelled and pre-processor list along with the clusters are utilized as an input for the purpose of evaluation of the output and the hidden layers of the artificial neural network. These selected attributes along with the clusters are essential for the implementation of the neurons that will be realized for the effective distribution of the weights as well as the bias weights.

$$Error\ Probability = \sum 1\ 2\ (T0\ -)\ 2 \text{ ----- } (2)$$

Where, T = Target Values

OL = Output Layer Values

Two target values are also assigned which are useful in identification of the hidden layer output through the input layers consisting of the attributes. The evaluation of the error probability is performed through the equation 2 given below. For the purpose of error probability estimation, the output layer values need to be evaluated. For this purpose, a ReLU activation function is deployed which achieves the target values of t1 and t2 with effective efficiency. These values are there provided for the calculation of error probability which is termed as inversely proportional to the precious of the probability achieve through the classification in the next step of this proposed approach. The process of Hidden layer estimation can be depicted in algorithm 1.

IV. Algorithm 1: Hidden Layer Estimation

```
//Input: Pre-processed List PPL, Weight set WSET = { }
//Output: Hidden Layer value list H
//Output: Hidden Layer value list HVL hiddenLayerEstimation (PPL, WSET), index=0
1: Start
2: HVL =∅ {Hidden Layer value]
3: for i=0 to size of PPL
4: ROW= PPL [i]
5: for j=0 to size of ROW
6: A=0
7: for k=0 to N [Number of Neurons]
8: ATR=ROW[j]
9: A = A + (ATR* WSET [index])
10: index++
11: end for
12: HVL = reLUmax (0, A)
13: end for
14: end for
15: return HVL
16: Stop
```

Step 4: Decision Making

The probability achieved in the previous step through the utilization of artificial neural network is essential for the realization of the yield prediction performed in this methodology. This step of the procedure takes the error probability list generated previously for the purpose of classification of the output for precise crop yield prediction and nutrient deficiency evaluation. The error probability list is first sorted in the ascending order and the top most values with the maximum error probability are utilized and segregated. The corresponding rose for these values are effectively extracted and correlated with the user input to achieve accurate crop yield prediction and fertilizer suggestion which is displayed to the user in the form of an interactive graphical user interface of the standalone application.

V. Result And Discussions

The proposed methodology for agricultural production estimation and fertilizer suggestion was developed using the Java programming platform. For the implementation of the suggested approach, the NetBeans IDE was employed. The installation device is equipped with 8 GB ram and 1TB of storage space, as well as an Intel Core i3 CPU. The JXL API is being used to allow the dataset to be interfaced with java program in a spreadsheet format.

An in-depth investigation of the prediction approach for the prevalence of any errors has been obtained through experimental examination. The error evaluation determines the reliability of the prediction, which may be particularly useful in determining the crop yield prediction system's efficiency. The error of the proposed prediction model was calculated using the RMSE (Root Mean Square Error) quality measurement approach.

The RMSE approach determines the error among two dependent and consistent parameters using association between two persistent components. The projected crop yield predictions and the observed crop yield predictions are the parameters in our technique. The error was calculated using the formula indicated in equation 3 below.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(x_{1,i} - x_{2,i}\right)^2}{n}} \quad ---(3)$$

Where, $\sum$ - Summation,$(x1 - x2)^2$ - Differences Squared for the summation in between the projected crop yield predictions and observed crop yield predictions, n - Number of samples or Trails

Table 1: Mean Square Error Measurement

No. of Trials	No. of Projected Crop Yield Predictions	No. of Observed Crop Yeild Predictions	MSE
1	11	10	1
2	5	4	1
3	9	9	0
4	10	8	4
5	2	2	0
6	1	1	0
7	6	4	4
8	14	12	4
9	3	3	0
10	12	10	4

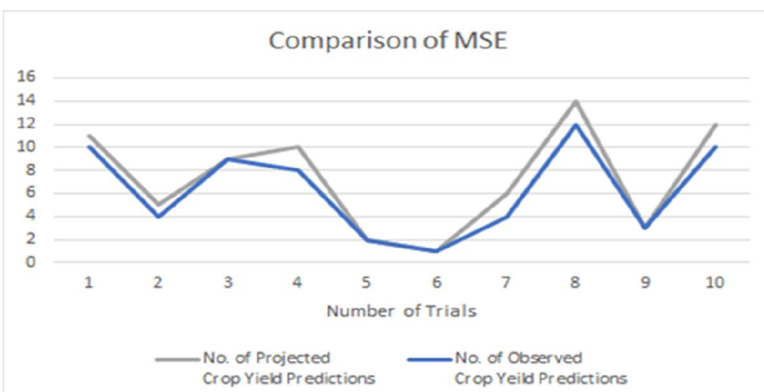

Figure 2: Comparison of MSE in between number of expected crop yield predictions V/s Number of obtained crop yield rate predictions

In the table 1 above, the experimental process and its outcomes are indicated. The data collected in the table is utilized to create a line chart, as illustrated in figure 2. We may conclude that the error produced in this approach for crop yield prediction is negligible after a thorough examination of the visually represented and tabulated results. To calculate the MSE, or Mean Square Error, a series of ten experimental trials were done with fluctuating input.

The results of the experiments show that the prediction system's error is appropriate and satisfactory. The prediction error is consistently observed in prediction models that make forecasts based on real-world input. The crop yield projections are influenced by a variety of distinct circumstances. The MSE and RMSE values of 1.8 and 1.34, correspondingly, are quite good and show that the crop yield forecast model was correctly implemented.

VI. Conclusion And Future Scope

The proposed methodology for agricultural production prediction and fertilizer suggestion using Artificial Neural Networks has been detailed in this research paper. As an input to the system, this method uses a dataset and user characteristics. The dataset and user input both contain essential crop statistics. This information is initially pre-processed to eliminate any redundant or contradictory information. This pre-processed data is sent on to the procedure's following stage, which involves labelling the data; the labelled data is then passed on to the procedure's following stage, which involves evaluating the clusters. For the generation of the clusters, which accomplishes the inner and outer clusters and sends the outcome to the following stage for neuron generation, linear clustering is utilized. To obtain the error probability in the attributes, the Artificial Neural Networks were developed employing the ReLU activation function. These variables are then utilized to create an output that is subjected to categorization using Decision Making in order to predict crop yields. The technique is evaluated for its effectiveness, and it receives a satisfactory outcome.

Future research might focus on integrating the crop yield prediction and fertilizer suggestion technique into the mobile application design, allowing any smartphone user to access agricultural yield projections.

References

[1]. S. Puengsungwan and K. Jirasereeamornkul, "Internet of Things (IoTs) based hydroponic lettuce farming with solar panels," 2019 International Conference on Power, Energy and Innovations (ICPEI), 2019, pp. 86-89, doi: 10.1109/ICPEI47862.2019.8944986.

[2]. D. Sunehra and M. Srinidhi, "Implementation of Smart Urban Farming using Raspberry Pi, Arduino and Node-RED Platform," 2020 IEEE International Conference for Innovation in Technology (INOCON), 2020, pp. 1-6, doi: 10.1109/INOCON50539.2020.9298357.

[3]. B. Ban, J. Lee, D. Ryu, M. Lee and T. D. Eom, "Nutrient Solution Management System for Smart Farms and Plant Factory," 2020 International Conference on Information and Communication Technology Convergence (ICTC), 2020, pp. 1537-1542, doi: 10.1109/ICTC49870.2020.9289192.

[4]. T. Xiong and S. Wang, "Intelligent farm management and control system based on Raspberry Pi," 2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2019, pp. 1286-1290, doi: 10.1109/IMCEC46724.2019.8983869.

[5]. N. Islam, B. Ray and F. Pasandideh, "IoT Based Smart Farming: Are the LPWAN Technologies Suitable for Remote Communication?" 2020 IEEE International Conference on Smart Internet of Things (SmartIoT), 2020, pp. 270-276, doi: 10.1109/SmartIoT49966.2020.00048.

[6]. S. R, V. G, K. M and L. J. K, "Intelligent Farming using Delta Robot," 2020 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), 2020, pp. 1-5, doi: 10.1109/ICPECTS49113.2020. 9337002.

[7]. N. Bore et al., "AGWS: Blockchain-enabled Small-scale Farm Digitization," 2020 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), 2020, pp. 1-9, doi: 10.1109/ICBC48266.2020.9169450.

[8]. H. chung, D. kim, S. lee and S. Cho, "Smart Farming Education Service based on u-learning environment," 2019 21st International Conference on Advanced Communication Technology (ICACT), 2019, pp. 471-474, doi: 10.23919/ICACT.2019.8701949.

[9]. H. P. Thakor and S. Iyer, "Development and Analysis of Smart Digi-farming Robust Model for Production Optimization in Agriculture," 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom), 2019, pp. 461-465.

[10]. X. Jiang et al., "Wireless Sensor Network Utilizing Flexible Nitrate Sensors for Smart Farming," 2019 IEEE SENSORS, 2019, pp. 1-4, doi: 10.1109/SENSORS43011.2019.8956915.