

# Performance Evaluation of Optimizing Crop Recommendation System in Machine Learning

BONAM ANUSAI SURYA KUMARI<sup>1</sup>, CH. RANJITH KUMAR<sup>2</sup>

<sup>1</sup>M. Tech Student, Dept. of CSE, BVC Institute of Technology and Science, Amalapuram, AP.

<sup>2</sup>Associate Professor & HOD (CSE-AI & DS, CSE- AI & ML), BVC Institute of Technology and Science, Amalapuram, AP.

**Abstract:** A large section of the Indian population considers agriculture as their primary occupation. Plant production plays an important role in our country. Low yields of good crops are due to excessive use of regular fertilizers and inadequate fertilizers. The proposed IoT and ML machine has been enabled for soil testing using sensors based entirely on measurement and observation of soil parameters. This system reduces the likelihood of soil degradation and allows crop health to be maintained. Different sensors with soil temperature, soil moisture, pH and NPK are used to monitor soil temperature, humidity and pH, and soil NPK vitamins, respectively. The data obtained by these sensors are stored in a microcontroller and analyzed using a system for learning algorithms such as Random Forest based on guidelines for the best crop are made. This paper also has a process that concentrates on using a convolutional neural network (CNN) as the main method of recognizing if the plant is at risk of a disease or not.

**Keywords:** Machine Learning, Convolutional Neural Network, Nitrogen-Phosphorus-Potassium, Crop Recommendation.

## I. INTRODUCTION

Agriculture is a vast region for the Indian economy and human survival. It is one of the basic occupations that are very important for human life. It also contributes greatly to our daily existence [1]. In most cases, farmers commit suicide due to reduced yields because they cannot repay the bank loans taken for farming activities [2]. We have seen in the current situation that the climate is constantly changing, which is harmful to plants and leads farmers to debt and suicide. It can mitigate by using various mathematical or statistical techniques on the information. With the help of these techniques, we can advise the farmer to get a satisfactory crop of his agricultural land to get maximum benefit. to gain. Today, agriculture in India is very advanced. "Website specific" farming is the key to proper farming. Although

health-related agriculture has improved, it still faces positive challenges. Healthy agriculture plays a key role in plant recommendation. The recommendation of plants is based on different parameters. Healthcare agriculture specializes in setting these parameters in site-specific ways to detect problems. Not all results provided by Health Agriculture are accurate down to the bottom line. In agriculture, however, it is important to have accurate and unique recommendations because, in the event of mistakes, this can result in a large amount of fabric and loss of capital. Much research is being finalized to obtain an accurate and more efficient model for crop forecasting.

In the proposed tool, environmental parameters, including rainfall, temperature, and state, are considered geographical areas and soil characteristics, including soil type, pH value, and vitamin concentration. Yes, to suggest a suitable crop. In addition to this, if the farmer chooses the right crop for the consumer, he will get a forecast of additional production. The aim is to develop a robust version of a particular country to provide accurate and accurate crop sustainability predictions for a specific type of land and climatic conditions, providing appropriate crop recommendations in this area so as not to harm the farmer further. Estimates of losses and income of various crops are mainly based on the previous year's statistics.

The proposed device is implemented through the study of systems, one of the applications of artificial intelligence that allows systems to be mechanically studied and developed without the help of a programmer. This will allow the accuracy of the program to move forward without human intervention. Many researchers are exploring this area to help farmers choose the appropriate crop in the selection below that considers various factors such as physical, environmental and financial factors.

Artificial neural networks are used to select crops with high yield rates [1], before planting, based on crop classification Decision Tree Learning-ID3 (Iterative Dichotomiser 3) and KNearest Neighbors Regression algorithms. It was done on Crop functions based on random forest algorithms, and BigML was analyzed. Machine learning algorithms are designed to prevent the effects of water pressure on plants and provide strict and fast selection principles used in predicting plant conditions. Machine learning techniques were used to calculate plants' value and intelligent systems to provide real-time suggestions. Several packages of gadget mastering algorithms in agricultural production frameworks have been surveyed in this work. It used other AI-powered systems to offer crop control suggestions. It can use deep learning techniques to get better yields in the crop. This article uses real-time month-to-

month seasons to design green production forecasting methods. A non-parametric statistical model was introduced with non-parametric regression techniques to implement the prediction method.

## II. LITERATURE SURVEY

**Kumar, Y. et al**[3] It can obtain crop production predictions of the proposed device from historical data, including temperature, humidity, pH, rainfall, and crop name. This system will preserve the maximum variety of crops in unique districts of India. With this proposed machine, we will predict the best crop according to the weather conditions of the field. This crop prediction can be accomplished using random forest algorithms and selection trees. An excellent accurate value is obtained using the random jungle set of rules. More accurate results gave more benefits to crop production.

**Suresh, G et al.** [4] This proposed tool is used to discover specific cultures according to the specific information provided. A support vector machine (SVM) has resulted in greater accuracy and productivity. This research paper mainly worked on two datasets: Place Statistics Patterns and Crop Statistics Sample Dataset. A crop was recommended according to its nutritional values (N, P, K, and PH) using this proposed tool and, in addition, the available nutritional values and the fertilizer required for the specific crop. Such as rice, corn, black gram, carrot and radish.

**Reddy, D et al.** [5] the proposed device worked on 3 parameters: soil characteristics, soil types and crop production record collection based primarily on the parameters that suggest the right crop for the farmer. The proposed device works on unique system learning algorithms such as Random Forest, CHAID, K-Nearest Neighbor, and Naïve Bayes. We can predict specific crops under specific climatic conditions and kingdom and district values using this proposed machine. Therefore, our proposed work will help farmers plant the right seed to increase the country's productivity, mainly based on the needs of the land.

**Rajak et al.** [6] This proposed technique is mainly used to determine specific crops based on a soil database. The proposed machine worked on various plants such as peanuts, beans, cotton, vegetables, bananas, rice, millet, sugarcane, coriander, and other properties such as

depth, texture, pH, soil colour, permeability, drainage, water maintenance and reduction. The proposed device worked on various device control classifiers such as Helper Vector Device (SVM), ANN, Random Forest, and Naïve Bayes. The term could be proposed for accurate site parameters with accuracy and performance. This studio painting can help farmers increase agricultural productivity, prevent soil degradation on cultivated land, reduce chemical use in crop preparation, and use environmentally friendly water resources.”

**Doshi et al. [7]** In this research work, they developed an intelligent system called Agro Consultant. The proposed system can be divided into two subsystems: i) crop forecasting and ii) precipitation forecasting. The proposed system consists of five essential items (millet, millet, maize, rice and wheat) and fifteen basic items (barley, cotton, peanuts, gram, jute, various beans, potatoes, ragi, tur, rapeseed and mustard, sesame, soybean,.) Works on. Sugarcane, sunflower, tobacco) Plants and some attributes include soil type, aquifer thickness, soil pH, topsoil thickness, rainfall, temperature, and location parameters. The proposed system implemented specific systems mastering algorithms such as the Decision Tree, K-nearest neighbour (K-NN), Random Forest, Neural network, and a multi-label class. The proposed system achieved an accuracy of 71% using the rain forecast model and 91.00% accuracy using neural network algorithms in the crop forecasting system.

**Kulkarni et al. [8]** The proposed system recommends the right crop based entirely on soil type and unique characteristics, such as average rainfall and soil temperature. The proposed machine worked on several machine learning algorithms, such as Random Forest, NaiveBase, and Liner SVM. This crop advisory machine-classified the input soil dataset into the recommended crop type, Kharif and Rabi. Using this proposed system, he completed 99.91% accuracy.

### III. PROPOSED MECHANISM

In our proposed investigation, each environmental and soil parameter has been considered. A certain form of soil will support the crop, while weather conditions will not guide it, thus affecting the yield. The universal operation of the proposed device is shown in Figure 1.

#### A. DATA ANALYSIS

This is an attempt to find out if there is any connection between the different attributes in the dataset.

### **Acquisition of Training Dataset:**

Various datasets are fed into the system from the official website [16] and Kegel [17]. The data set consists of i) Production Dataset: This dataset includes the production of 16 major crops grown per kilogram per hectare in all the states. Zero yields indicate that the crop is not grown in the country concerned. ii) Crop cost data set: This record set gives the price of each crop in rupees. By hectare iii) Modal rate of vegetables: This dataset provides general market prices of this vegetable for two months. iv) Standard Price of Vegetables: This dataset provides the current market rate of crops at Rs. Per hectare. V) Soil Nutrient Dataset: This dataset consists of five columns with attributes in the order of status, nitrogen content, phosphorus content, potassium content and general pH. Vi) Rainfall temperature data set: This data set includes plants, maximum and minimum rainfall, maximum and minimum temperature, maximum and minimum rainfall and pH values. Revenue analysis is conducted using crop price, market charge, current price and production data set. It is being done as a first step in understanding how earnings will significantly impact crop forecasting. Profit is calculated for each crop grown within the country and sets a value of -1 for states where a given crop has 0 or no production.

### **DATA PREPROCESSING**

This step involves changing the null and 0 values for the output using -1 so that it does not further affect the general prediction. Further, the set of data must be encoded to be fed into the neural network. Data pre-processing is important because it allows records to be refined and adapted for use in device learning algorithms. The focus in pre-processing is on eliminating any lost or misplaced records and dealing with any lost values. The values in the dataset are in string format. It must be converted to numeric values to leave input in the neural community. In addition, to reduce the number of records in the linear regression model, plants are being filtered primarily based on the gift of specific vitamins and nutrients in the soil. If the nutrients in the soil are less than the required amount due to the use of crops, then the relevant crop can be wasted, in which school time has been greatly reduced.

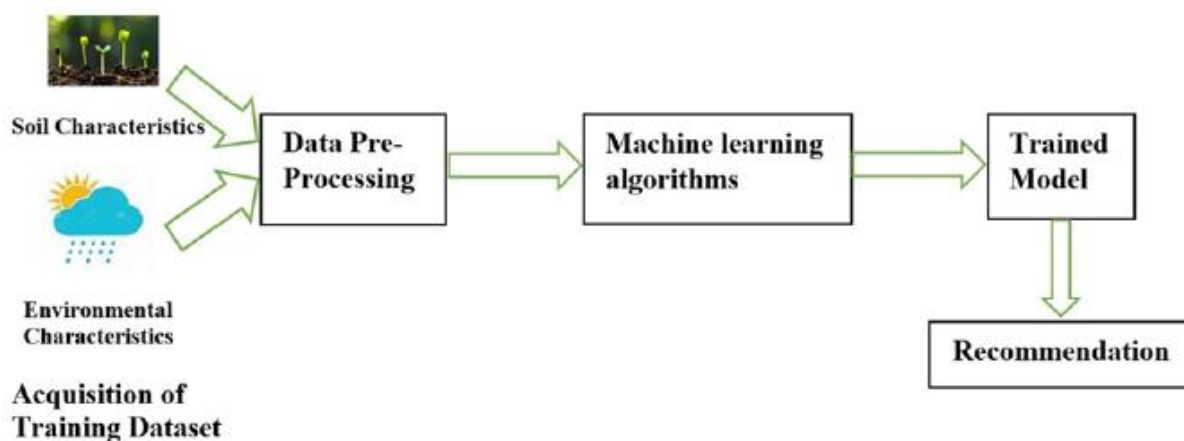
### **DECISION TREE**

Decision tree classifiers use greedy techniques. It is a set of supervised learning rules in which a tree represents class attributes and labels. The most important purpose of using the Decision Tree is to form a training prototype that we can use to predict the magnificence or

cost of the target variables through learning selection policies deduced from previous facts (training statistics). The decision tree can be described in two wonderful types: selection nodes and leaves. The leaves are the results or the last parry effects. Each node within the tree acts as a check case for some feature, and each area descending from that node corresponds to one of the check case's feasible responses. This procedure is recursive in nature and is repeated for each subtree rooted at the new nodes.

**LINEAR REGRESSION**

Linear regression fits a straight line between rainfall, temperature, pH, and yield, which can be regressed to the predicted value and for each crop. At the dropout, the crops are sorted by linear regression version based on your price. Short-level use yields a first-level score on the list.



**Fig.1** Crop Recommendation System

**Random Forest (RF)**

Random Forest is a set of ML rules. In the case of education, more than one selection is traced, and the output will be divided based on the training range, i.e. type, class prediction, i.e. regression. The number of bushes is proportional to the accuracy of the forecast. The data set includes rainfall, perception, temperature, and yield elements. These factors in the dataset are used for schooling. Only two-thirds of the data set is considered. The rest of the dataset is used for experimental purposes. The Random Woodland algorithm has 3 parameters, such as the n tree, that defines the n range of shrubs that need to be grown m Effort: Specifies the number of variables that need to be taken into the node split. Node Size: In terminal nodes, it indicates the number of observations we should receive.

**XGBoost**

"Extreme Gradient Boosting (XGBoost) is an innovative, adaptive interpretation of the Gradient Boosting algorithm designed for release feasibility, computational speed, and performance. It is an open-source library and is part of the distributed machine learning community. XGBoost offers a parallel tree boost (also called GBDT, GBM) that can speed up and correct many problems of registry technology."

**IV. EXPERIMENTAL RESULTS**

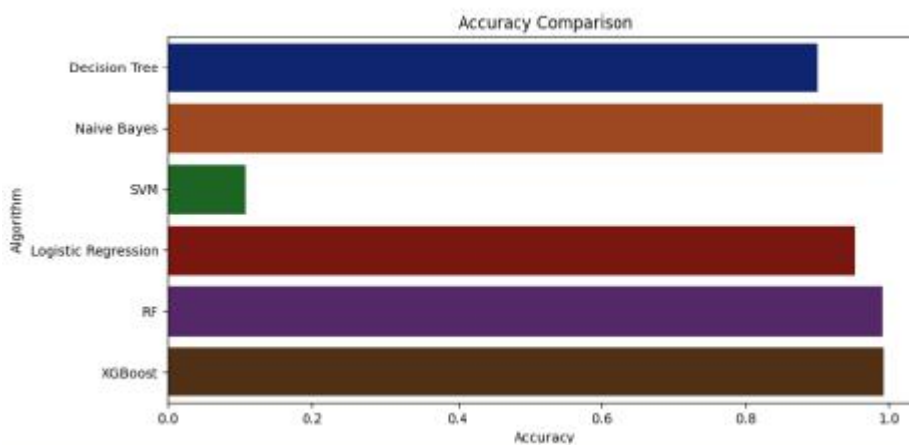


Fig.2 Accuracy Comparison

```

Decision Tree --> 0.9
Naive Bayes --> 0.990909090909091
SVM --> 0.106818181818181
Logistic Regression --> 0.9522727272727273
RF --> 0.990909090909091
XGBoost --> 0.9931818181818182
    
```

Fig.3 Machine Learning vice Accuracy Result

**V. CONCLUSION**

This paper successfully proposed and implemented an intelligent crop recommendation system that farmers across India can easily use. This machine will help farmers make informed choices based on certain parameters such as nitrogen, phosphorus, potassium, pH value, humidity, temperature and rainfall. Using this research, we will increase the country's

productivity and transfer revenue in this way. In this way, the farmer can increase his yield by cultivating the right crop and increasing the UK's gross profit. This research has advised various plants in India to use specific device study algorithms like Decision Tree, Naïve Bayes, Support Vector Machine, Logistic Regression, Random Forest and XGBoost. This analysis was performed on these six variants of the device control algorithm, and out of these six algorithms, XGBoost completed a very accurate result.

## REFERENCES

- [1] R. Kumar, 2015, "Crop Select ion Method to maximize crop yield rate using machine learning technique", pp. 138-145
- [2] H. Lee and A. Moon, 2014, "Development of yield predict ion system based on real-t ime agricultural meteorological information", pp. 1292- 1295.
- [3] T.R. Lekhaa, 2016, "Efficient Crop Yield and Pesticide Predict ion for Improving Agricultural Economy using Data Mining Techniques".
- [4] Kumar et al. 2020, "Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector.", pp. 736-741.
- [5] Nigam, 2019, "Crop yield prediction using machine learning algorithms", pp. 125-130, IEEE.
- [6] Medar, and Shweta Shweta, 2019, "Crop yield prediction using machine learning techniques.", pp. 1-5. IEEE
- [7] . Jain, Sonal, and Dharavath Ramesh. 2020, "Machine Learning convergence for weather based crop selection.", pp. 1-6. IEEE.
- [8] . Gandge, Yogesh, 2017, "A study on various data mining techniques for crop yield prediction.", pp. 420-423.
- [9] . Suresh and M. Ramalatha, 2018, "Prediction of major crop yields of Tamilnadu using K-means and Modified KNN", pp. 88-93. IEEE