Machine Learning Based Crop Recommendation System Using Convolutional Neural Network

K VISWANATH, VIJAYA BHASKAR MADGULA, DIGALA RAGHAVARAJU
ASSISTANT PROFESSOR [1,2,3]
viswanath.k003@gmail.com, vijaya.bhaskar2010@gmail.com, raghava.digala@gmail.com
Dept.of CSE, Sri Venkateswara Institute of Technology, N.H 44, Hampapuram, Rapthadu, Anantapuramu, Andhra Pradesh 515722

Abstract: A sizable portion of India's population views farming as their main source of income. An essential part of our nation's economy is the cultivation of plants. Both insufficient and over-application of conventional fertilisers lead to poor crop yields. Soil testing using sensors that depend only on measuring and observing soil properties is now possible with the help of the suggested IoT and ML machine. Soil deterioration is less likely to occur and crop health may be preserved using this technique. Soil temperature, humidity, pH, and NPK vitamin monitoring is done by use of separate sensors for each of these variables. A microcontroller stores the sensor data, which is then processed using a learning algorithm system, such as Random Forest, to determine the optimal crop. In addition, this research details a procedure that primarily use a convolutional neural network (CNN) to determine if the plant is susceptible to a disease.

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

I. INTRODUCTION

A large portion of India's economy and human existence depend on agriculture. For human survival, it is one of the most fundamental jobs. It is also an essential part of our everyday lives [1]. Reduced harvests make it impossible for farmers to repay the debts they took out from banks, and as a result, many farmers take their own lives.

things to do. Climate change, as we have seen in the present, is destructive to plants and causes farmers to go into debt or even commit suicide. Applying different statistical or mathematical methods to the data might help reduce it. We can guide the farmer towards a successful harvest by using these methods.

of his farmland to reap the most possible rewards. Agriculture in India has come a long way in recent decades. To farm properly, one must do "website specific" farming. Agriculture with a focus on health has come a long way, but there are still obstacles to overcome. Plant advice is an important part of healthy agriculture.Various factors are considered while making plant recommendations. In healthcare agriculture, these characteristics are put up in a manner

that is particular to each location in order to identify issues. When it comes down to it, not all of Health Agriculture's findings are accurate. But, in agriculture, precise and original advice is crucial, since blunders may lead to significant fabric waste and financial losses. A more precise and effective model for crop forecasting is the goal of ongoing research.

Soil type, pH value, vitamin content, and environmental aspects like rainfall, temperature, and condition are all taken into account by the suggested technique. To recommend an appropriate crop, yes. Beyond this, the farmer may anticipate increased output if he selects the appropriate crop for the buyer. Creating a reliable model of a certain nation that can reliably forecast the long-term viability of its crops is the ultimate goal.

a particular kind of terrain and climate, advising farmers on what crops might thrive there without causing any further damage. The majority of the revenue and loss estimates for different crops are derived on the data collected in the prior year.

The suggested gadget is put into action by use of system analysis, a branch of AI that permits the mechanical study and development of systems without the intervention of a programmer. Because of this, the programme will be able to continue accurately even when no humans are present. Numerous researchers are delving into this field to assist farmers in selecting the most suitable crop from the options provided, taking into account aspects like physical, environmental, and economical considerations.

In order to choose crops with high yield rates, artificial neural networks are trained using methods like K-Nearest Neighbours Regression and Decision Tree Learning-ID3 (Iterative Dichotomiser 3). This is done before planting. Crop functions derived from random forest techniques were used for this analysis, which was conducted on BigML. To forecast plant conditions and protect them from water pressure, machine learning algorithms provide rapid and rigorous selection criteria. Artificial intelligence

methods were used to determine the monetary worth of plants, and smart systems were employed to provide ideas in real-time. This article surveys several packages of agricultural production frameworks that use gadget mastering methods. In order to provide recommendations for crop management, it drew on other AI-powered systems. It is capable of improving agricultural yields with the use of deep learning methods. This article presents approaches for green production forecasts that employ real-time month-to-month seasons. The prediction approach was implemented using a non-parametric statistical model that made use of non-parametric regression methods.

II. LITERATURE SURVEY

Kumar, Y. et al[3] By analysing past data such as temperature, humidity, pH, rainfall, and crop name, it is possible to derive projections of the proposed device's crop yield. The greatest possible diversity of crop types in distinct Indian areas will be preserved by this strategy. This suggested machine will analyse the field's meteorological data and provide a crop prediction based on it. Using selection trees and random forest methods, this crop prediction may be achieved. The random jungle rules provide a very precise result. Crop production reaped greater advantages from more precise findings.

Suresh, G et al. [4] Using the supplied data, this suggested technology may identify certain civilizations. Performance and accuracy have both been enhanced by using a support vector machine (SVM). Two datasets, Place Statistics Patterns and Crop Statistics Sample Dataset, were the primary foci of this study article. This suggested technique was used to recommend crops based on their nutritional values (N, P, K, and PH), accessible nutritional values, and the fertiliser needed for that particular crop. Things like black gramme, maize, carrots and radishes are examples.

Reddy, D et al. [5] The suggested tool relied on three factors to determine the best crop for farmers: soil type, soil features, and the ability to record crop productivity. The suggested gadget is compatible with distinct system learning algorithms including Naïve Bayes, Random Forest, CHAID, and K-Nearest Neighbour. This suggested computer can forecast the worth of kingdoms and districts as well as individual crops in response to unique weather patterns. Consequently, our planned project will provide farmers the tools they need to boost national output by planting seeds tailored to certain soil types.

Rajak et al. [6] The primary use of this suggested method is the identification of certain crops. using data from a soil library. Peanuts, beans, cotton, veggies, bananas, rice, millet, sugarcane, coriander, and other plant characteristics including depth, texture, pH, soil colour, drainage, permeability, water management, and reduction were all

successfully processed by the suggested equipment. The device that was suggested might be used with many types of device control classifiers, including SVM, ANN, Random Forest, and Naïve Bayes. For precise and effective site characteristics, the word might be suggested. Improved agricultural output, less chemical usage during crop preparation, less water waste, and less soil deterioration on farmed land are all possible outcomes of this studio painting.

Doshi et al. [7] An AI system named Agro Consultant was created as a result of this study. Crop forecasting and precipitation forecasting are the two main components of the suggested system. Five staples—millet, millet, maize, rice, and wheat—and fifteen non-essentials—barley, cotton, peanuts, gramme, jute, different beans, potatoes, ragi, tur, rapeseed, mustard, sesame, and soybean—make up the suggested system. Cane sugar, sunflower oil, and tobacco Plants and a number of characteristics, such as soil type, aquifer depth, soil pH, topsoil thickness, climate, rainfall, and

parameters for the place. Decision trees, K-nearest neighbours (K-NN), random forests, neural networks, and multi-label classes were all included into the proposed system. By using neural network algorithms in the crop forecasting system, the suggested system was able to get an accuracy of 91.00% and a rainfall prediction model accuracy of 71%.

Kulkarni et al. [8] Soil temperature, average rainfall, and type are the only variables that the suggested approach takes into account when making crop recommendations. Random Forest, Linear Support Vector Machine, and NaiveBase were among the machine learning techniques that the suggested machine may use. Based on the input soil information, this crop advice computer advised either Kharif or Rabi crops. He achieved an accuracy rate of 99.91 percent when using the suggested approach.

PROPOSED MECHANISM

All soil and environmental factors have been taken into account in our planned study. The crop's support system is the soil type, and the yield is affected by the weather conditions, which do not direct the crop. In Figure 1, we can see how the suggested gadget works in every situation.

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:

Both the official website [16] and Kegel [17] provide datasets that are input into the system. Part of the collection is the "Production Dataset," which compiles state-by-state statistics on the yield of sixteen different crops in kilogrammes per hectare. When crop yields are zero, it means the crop is not cultivated in that particular nation. ii) A dataset including information on agricultural costs: this database lists the prices of various crops in rupees. In terms of hectares iii) Vegetable modal rate: This dataset gives two months' worth of average market pricing for this vegetable. iv) Vegetable Standard Price: This information gives the average market price of crops per hectare in Indian rupees. There are five columns in the Soil Nutrient Dataset, and the properties are listed as follows: status, nitrogen content, phosphorus content, potassium content, and overall pH. Chapter VI: The Rainfall Temperature Data Set This dataset contains information on plants, as well as the highest and lowest recorded rainfall amounts, temperatures, pH values, and temperatures. Crop price, market charge, current price, and production data sets are used in revenue analysis. The process is being completed as a first stage in comprehending the substantial influence that profits will have on crop prediction. For each crop produced in the nation, a profit is determined; in states where a particular crop produces no harvest, the profit is set to -1.

DATA PREPROCESSING

To ensure that this step does not impact the overall forecast any more, we will use -1 to change the null and 0 values for the output. In order to feed the data set into the neural network, it must also be encoded. Records may be fine-tuned and adjusted for use in device learning algorithms by data pre-processing. The primary goal of pre-processing is to deal with missing data and remove any entries that have been misplaced. The collection contains values that are formatted as strings. For input to be left with the neural community, it must be transformed to numerical values. Moreover, plants are mostly screened according to the donation of certain soil elements and vitamins in order to decrease the number of records in the linear regression model. The time spent in school has been significantly reduced since crops might cause soil nutrients to be insufficient, leading to the waste of the relevant crop.

DECISION TREE

The greedy approaches are used by decision tree classifiers. A tree represents the properties and labels of classes in this collection of supervised learning rules. Predicting the value or cost of the target variables using learning selection rules derived from training data is the primary goal of using the Decision Tree as a training prototype. Two fantastic ways to characterise a decision tree are as leaves and selection nodes. The final impacts of parrying are shown by the leaves. At each node in the tree, there is a feature that has to be checked off, and the areas that descend from that node represent possible answers to that features. Each child tree that takes its cues from the newly added nodes undergoes the same cyclical process.

LINEAR REGRESSION

The projected value for each crop may be determined using linear regression, which fits a straight line between rainfall, temperature, pH, and yield. The crops are sorted at the dropout using a linear regression version that takes your pricing into account. A first-level score is achieved by short-level usage.

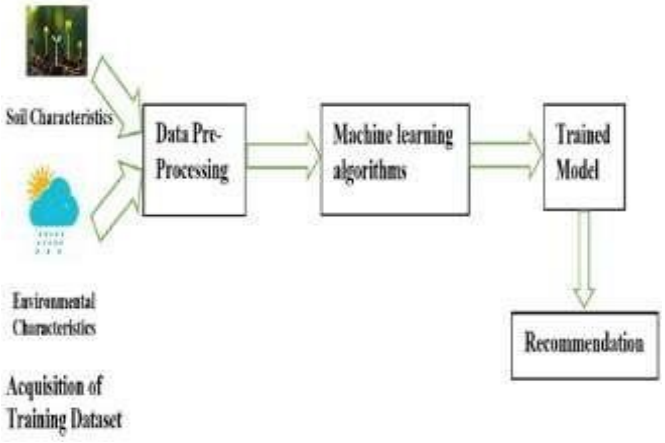

Fig.1 Crop Recommendation System

Random Forest (RF)

The Random Forest is a collection of rules for machine learning. Education involves tracing many selections and dividing the output according to the training range (type, class prediction, regression). The accuracy of the prediction is directly correlated to the quantity of bushes. Elements such as yield, temperature, perception, and rainfall are included of the dataset. The educational system makes advantage of these key factors. The data set is only examined up to the second half. Experiments make use of the remaining dataset. Three parameters make up the Random Woodland method; one of them is the n-tree, which specifies the n-range of shrubs that must be developed. The amount of variables that must be considered in the node split is specified by the effort parameter. Terminal nodes' node size specifies the expected number of observations.

XGBoost

With an eye on release feasibility, computing speed, and performance, Extreme Gradient Boosting (XGBoost) offers a fresh take on the Gradient Boosting method via adaptive design. It is a component of the distributed machine learning community and is an open-source library. In order to expedite and resolve several issues with registry technology, XGBoost provides a parallel tree boost, which is also known as GBDT or GBM.

III. EXPERMENTAL RESULTS

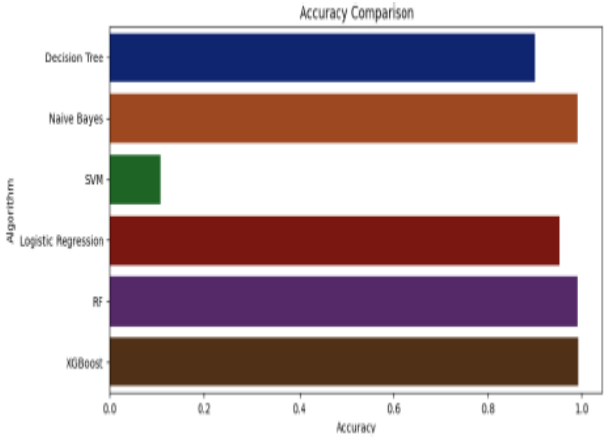

Fig.2 Accuracy Comparison

```
Decision Tree --> 0.9
Naive Bayes --> 0.990909090909091
SVM --> 0.10681818181818181
Logistic Regression --> 0.9522727272727273
RF --> 0.990909090909091
XGBoost --> 0.9931818181818182
```

Fig.3 Machine Learning vice Accuracy Result

IV. CONCLUSION

An intelligent crop suggestion system was developed and put into action in this article, which farmers all across India may use with ease. Based on factors like pH value, nitrogen, phosphorus, potassium, humidity, temperature, and rainfall, this equipment will assist farmers in making educated decisions. via the use of this study, we want to enhance the nation's production and generate more income via transfers. By growing the correct crop, the farmer may boost his output and the UK's gross profit. Different factories in India have been suggested by this research to employ certain device study techniques such as Decision Tree, Naïve Bayes, Support Vector Machine, Logistic Regression, Random Forest, and XGBoost. From these six variations of the device control method, XGBoost produced the most accurate result in the study.

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