International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

Prediction of Crop Based on Soil Properties

S. Keerthi, T. Esther Rani, B. Vivek, V. Sasidhar Reddy

UG Student,Assistant Professor, Department of Computer Science and Engineering

Kallam Haranadha Reddy Institute of Technology, Chowdavaram, Guntur, Andhra Pradesh, India

ABSTRACT

One of India's most important and prevalent professions, agriculture plays a crucial part in the growth of our nation. Improving crop output is therefore viewed as a significant aspect of agriculture since 60 percent of the nation's territory is used for agriculture to feed its 1.2 billion inhabitants. Basically, we need to know what kind of crop can be grown here if we have a plot of ground. The different aspects of dirt are important to agriculture. Crop production is a challenging job because it involves a variety of variables, including soil type, temperature, humidity, etc. However, due to unnatural climatic changes, food output and forecasting are currently declining, which will have a negative impact on farmers' economies by resulting in a low yield and also make farmers less adept at predicting future crops. Farmers and the other parties involved would benefit greatly from being able to locate the crop before sowing it in order to make informed choices regarding storage and business operations. By keeping track of the agricultural area based on the properties of the soil and advising farmers on the best crop to grow, the proposed project would help them to significantly boost output and lower loss. In our research, we develop a recommendation system that uses machine learning methods like logistic regression, support vector machines, and others to suggest the best crops based on the input soil parameters. Here, the seed information for the crops is gathered along with the necessary conditions, such as temperature, humidity, and moisture content, which aids in the crops' successful development. Thus, this method lessens the financial losses that farmers experience as a result of planting the incorrect crops. It also aids farmers in discovering new crop varieties that can be grown in their region.

Keywords: Machine Learning, Crop Prediction, Soil Properties.

1. INTRODUCTION

In Industry that is 4.0, also known as the Fourth Industrial revolution, The primary driver of the Indian economy is agriculture. Agriculture has long been regarded as one of the primary activities carried out in India. In India, agriculture employs 50% of the labour population. In terms of a few commodities, India is the top producer. India's primary industry is agriculture. The primary and fundamental component of agriculture is the soil. However, producers are still employing the old technique today. Farmers' inability to obtain satisfactory results using the conventional technique indicates that crop production is not growing. Good soil quality is necessary to boost crop yields. Crop quality and output are entirely dependent on the soil. The soil qualities used in agriculture include those linked to organic matter, such as nitrogen, phosphorus, and potassium (Potassium). We were inspired to create this method in order to assist farmers in selecting the crop that should be grown for their benefit. The dataset includes the nutrients that are readily accessible to farmers' soil, including N, P, K, humidity, rainfall, pH, and temperature. The crop that can grow in a specific soil is predicted

using a method that takes into account the soil type. As a result, the crop yield rises and the farmer makes more money using this novel technique. We use cutting-edge technology to build the system. The technology is built using machine learning. Machine learning focuses on developing software applications that can acquire data and use that data to learn. Machine learning enables the creation of models from sample data as well as the ability to make decisions autonomously based on prior knowledge.

1.1 Prediction of Crop

We were inspired to create this method in order to assist farmers in selecting the crop to plough for their benefit. The accessible nutrients for farmer's soil make up the dataset. Our system forecasts crop based on the value of the nutrients. The soil type method makes a prediction about the types of crops that will thrive in a given soil. As a result, the crop yield rises and the farmer makes more money using this novel technique. The technology is built using machine learning

2. LITERATURE SURVEY

The study explores exploratory data and takes into account various predictive model designs. Different regression techniques are attempted to identify and examine each property using a data set as a sample data set. To determine the best crop to cultivate, various algorithms were applied to the data collection, including K Nearest Neighbors, Naive Bayes, and KNN with Cross Validation[1].

The system that was created suggested the crop that would grow best on a specific plot of ground. based on soil composition and environmental variables like rainfall, temperature, humidity, and pH. To find patterns in the input data and handle it in accordance with the input requirements, Support Vector Machine (SVM) and Decision Tree machine learning predictive algorithms are used. The system suggested a crop for the farmer as well as how much fertilizer should be added for the anticipated produce. Other requirements for the system included displaying the estimated yield in q/acre, the amount of seed needed for cultivation in kg/acre, and the crop's market price [2].

This paper offers a method for smart agriculture through field monitoring, which can greatly help farmers increase output. In order to find patterns in the data and then process it in accordance with the input circumstances, it also uses machine learning and prediction algorithms like multiple linear regression. [3]

This study paper's objective is to suggest and put into practise a rule-based system to forecast crop yield production using historical data. By using association rule mining on agricultural data from 2000 to 2012, this was accomplished [4].

The project's main goal is to develop a prediction model that can be used to foretell the crop's maximum output rate before it is sown. A machine learning algorithm is applied to the data to estimate the output rate of crops based on the farmer's state, district, season, land area, and crop type [5].

Based on the literature survey 60% of India's territory is used for agriculture in order to feed its 1.3 billion inhabitants. Additionally, the populace is growing daily. Therefore, agriculture must be modernised in order to benefit producers in our nation and address many of their issues.

Farmers in the current setup have no access to technology or analysis. Farmers in the traditional system employ the "trial and error" technique. A farmer experiments on land with various crops,

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

water availability, etc., and after numerous such "tries," the farmer probably achieves the anticipated crop output.

Numerous papers have conducted the poll while taking into account various factors. There are some methods that aid in crop selection, but no system is perfect.

In some papers, crop yield predictions based on climatic input parameters are made using data mining methods. However, predicting crop output solely on the basis of climatic factors is insufficient.

Different machine learning algorithms that can be used to predict crops have been analysed in some survey studies.

There are numerous review articles on crop prediction that outline various prediction algorithms. But at this time, there isn't such a method. Therefore, it is necessary to put such a system in place so that farmers can profit from it.

3. PROPOSED SYSTEM

Implementation methodology
Implementation Steps:

- **Data Collection**
- **Data Pre-processing**
- **Training and Testing Data**
- **Result and analysis**

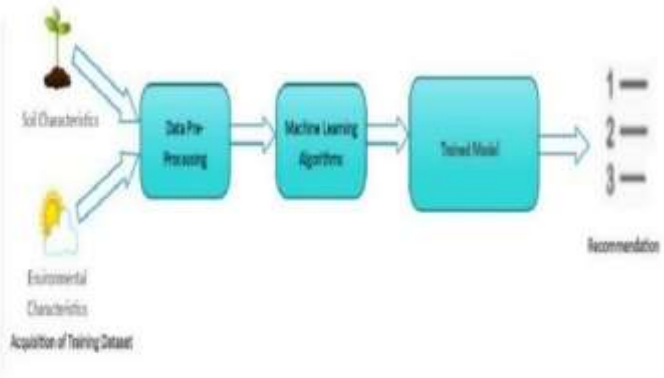

Fig-1: Proposed approach

3.1 Data collection:

One of the initial steps we perform during deployment is a data analysis. We carried out this analysis to check for correlations between the different dataset characteristics. Any machine learning method's accuracy is determined by the quantity of factors and the validity of the training dataset. This study meticulously selected the settings that would yield the best results after examining a variety of datasets from the Kaggle website. Environmental factors have been used in many studies on this topic to forecast crop sustainability; some have focused primarily on yield, while others have only considered fiscal factors. In order to provide the farmer with an exact and reliable recommendation on which crop would be best for his property, we combined climatic factors like rainfall, temperature,

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

and soil ph with soil parameters like soil nutrients. Using the read csv() function from the pandas package, we import the dataset.

H1 label

	A	B	C	D	E	F	G	H
1	N	P	K	temperati	humidity	ph	rainfall	label
2	90	42	43	20.87974	82.00274	6.502985	202.9355	rice
3	85	58	41	21.77046	80.31964	7.038096	226.6555	rice
4	60	55	44	23.00446	82.32076	7.840207	263.9642	rice
5	74	35	40	26.4911	80.15836	6.980401	242.864	rice
6	78	42	42	20.13017	81.60487	7.628473	262.7173	rice
7	69	37	42	23.05805	83.37012	7.073454	251.055	rice
8	69	55	38	22.70884	82.63941	5.700806	271.3249	rice
9	94	53	40	20.27774	82.89409	5.718627	241.9742	rice
10	89	54	38	24.51588	83.53522	6.685346	230.4462	rice
11	68	58	38	23.22397	83.03323	6.336254	221.2092	rice
12	91	53	40	26.52724	81.41754	5.386168	264.6149	rice
13	90	46	42	23.97898	81.45062	7.502834	250.0832	rice
14	78	58	44	26.8008	80.88685	5.108682	284.4365	rice
15	93	56	36	24.01498	82.05687	6.984354	185.2773	rice
16	94	50	37	25.66585	80.66385	6.94802	209.587	rice
17	60	48	39	24.28209	80.30026	7.042299	231.0863	rice
18	85	38	41	21.58712	82.78837	6.249051	276.6552	rice
19	91	35	39	23.79392	80.41818	6.97086	206.2612	rice
20	77	38	36	21.86525	80.1923	5.953933	224.555	rice
21	88	35	40	23.57944	83.5876	5.853932	291.2987	rice
22	89	45	36	21.32504	80.47476	6.442475	185.4975	rice
23	76	40	43	25.15746	83.11713	5.070176	231.3843	rice
24	67	59	41	21.94767	80.97384	6.012633	213.3561	rice
25	83	41	43	21.05254	82.6784	6.254028	233.1076	rice

Fig-2: dataset of crop recommendation System

3.2 Data Preprocessing:

Sometimes, real-world data has noise, missing values, and is in an unsuitable format that prevents it from being immediately incorporated into machine learning models. To clean data and make it suitable for a machine learning model, which increases the model's efficacy and accuracy, data preprocessing is a necessary job. Data cleaning and preparation for use in machine learning algorithms make data preprocessing a crucial stage. Preprocessing is primarily concerned with resolving any missing data as well as removing any outliers or inaccurate data. There are two methods to fill in any gaps in the data. The first choice is to remove the complete row that contains the inaccurate or missing data. Although this technique is straightforward to use, it works best with sizable datasets.

3.3 Training and Testing Data:

We used numerous ml methods to obtain accurate findings because the proposed model needs to be trained and tested in a variety of scenarios. Here, we've trained the data so that it can forecast the crop that can be grown based on a variety of provided parameters, such as environmental variables and soil nutrients. We train the data to forecast the precise crop to be grown using a variety of input parameters. We make forecasts based on the X test data and fit the data to the X, Y training values. We ran 100 training epochs on the model. The best model is one that has the lowest loss, and this model is used for testing and assessment.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

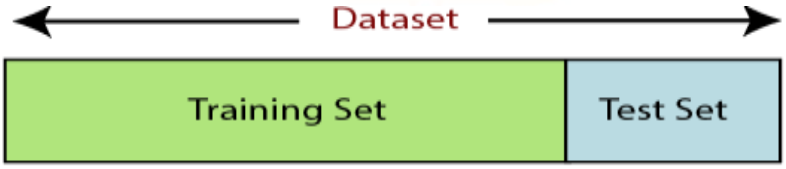

Fig-3: Splitting of dataset.

3.4 Result and analysis:

The forecast outcomes are evaluated using the accuracy parameter.

Accuracy:

When true positive and true negative are multiplied by a percentage of true positive, true negative, and false positive with false negative, the result is an estimate of how close the computation is to the actual value.

4. ALGORITHMS

classification using random forest
Journal of Engineering Sciences: Algorithms
1. We essentially pick the k to feature at random from all m features in the model.
2. Using the best split point, we compute the node d by choosing the k feature.
3. As a result, we split the nodes into daughter nodes using the split technique.
4. Continue doing steps 1 through 3 until you have the required number of nodes.
5. To make an endless number of trees, carry out steps 1-4 an infinite number of times to make a forest. to forecast using the learned random forest algorithm.

The method uses the following pseudo code:
1. To predict the output and the result, which were then saved, we used the test characteristics and each random decision tree.
2. After that, we calculated the vote that each decision tree offers for each result that is predicted.
3. Lastly, we examined the most widely anticipated result, which provides the ultimate prediction from the random forest.

K-Nearest Neighbour (KNN) algorithm:

1. The K-Nearest Neighbour (KNN) algorithm belongs to the class of supervised learning techniques and is one of the simplest machine learning methods.
2. The K-NN algorithm saves all the information that is accessible and categorizes new data based on similarity.
3. This means that using the K-NN algorithm, new data can often be quickly and accurately classified into a suitable group.
4. K-NN algorithms are frequently used for classification issues. o Regression issues are also addressed in some instances.

Decision Tree Algorithm:

1. The supervised learning algorithms group includes Decision Tree. The majority of classification and regression issues are solved using decision tree algorithms.

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

2. Each leaf node of the decision tree correlates to a class label, and the internal nodes of the tree are used to represent the attributes in order to solve the issue.
3. A decision tree only accepts yes or no as its only two Binary values.
4. If the response is yes, the tree is divided into another sub-tree; otherwise, the process halts and the node turns into a leaf node.

Support Vector Machine (SVM):

A supervised machine learning method is called the Support Vector Machine (SVM). Although the Support Vector Machine (SVM) is used for both categorization and regression, it is primarily used for classification. As a result of the high accuracy rate offered by the Support Vector Machine (SVM) method, we also used it. In this method, each piece of data is represented as a point in an N-dimensional area, and a hyper plane is built to divide the points into various classes. The hyper plane is then used to perform classification. The datasets will be divided into various classes as positive and negative by the hyper plane.

5. RESULTS

Fig-4: Home Page

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

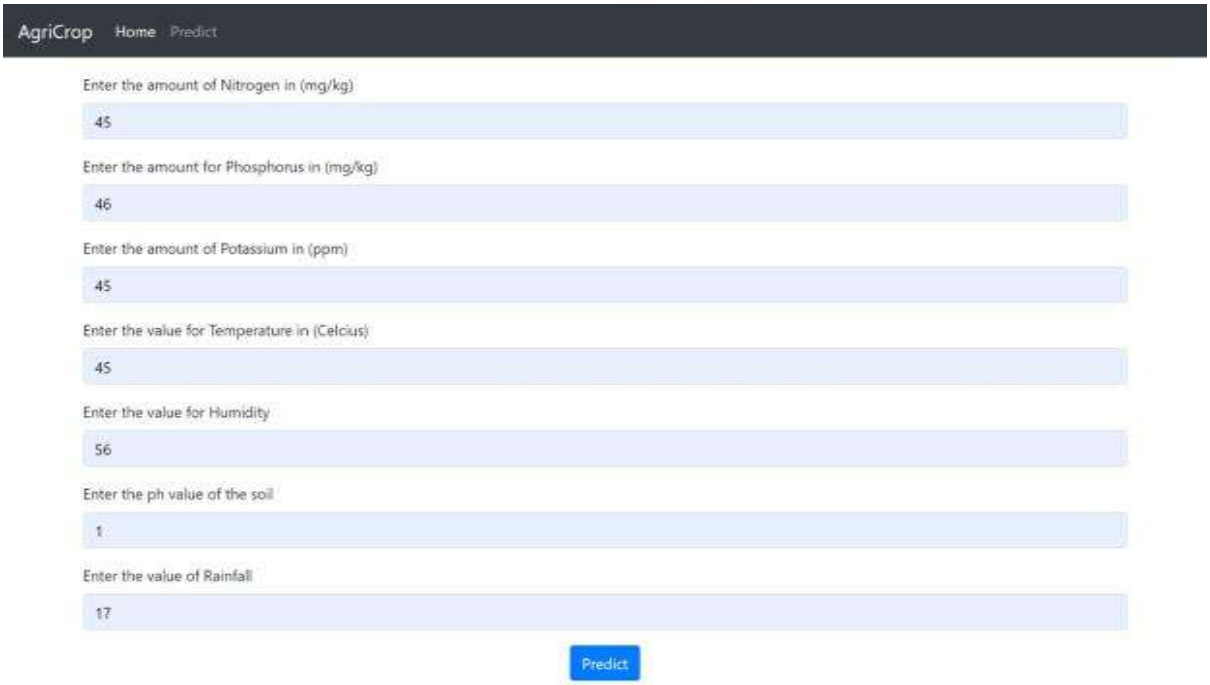

Fig-5: Soil details form

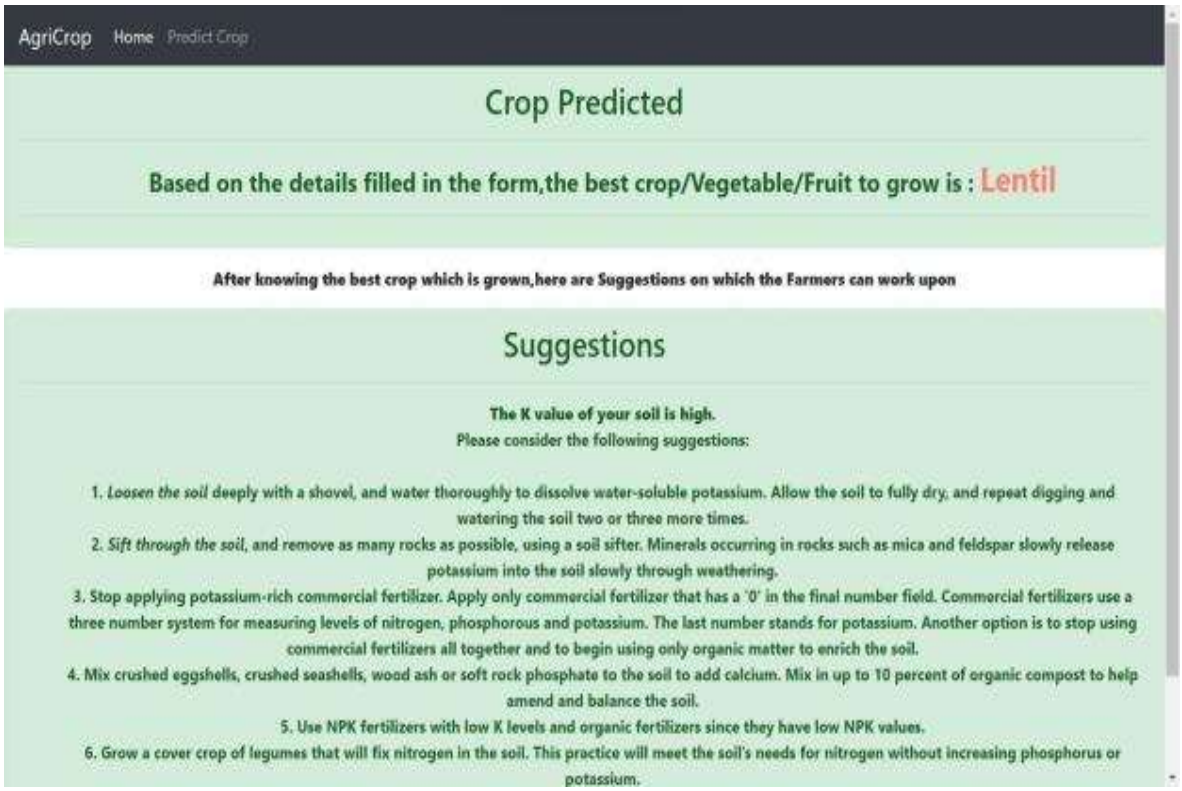

Fig-6: Crop Predicted

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

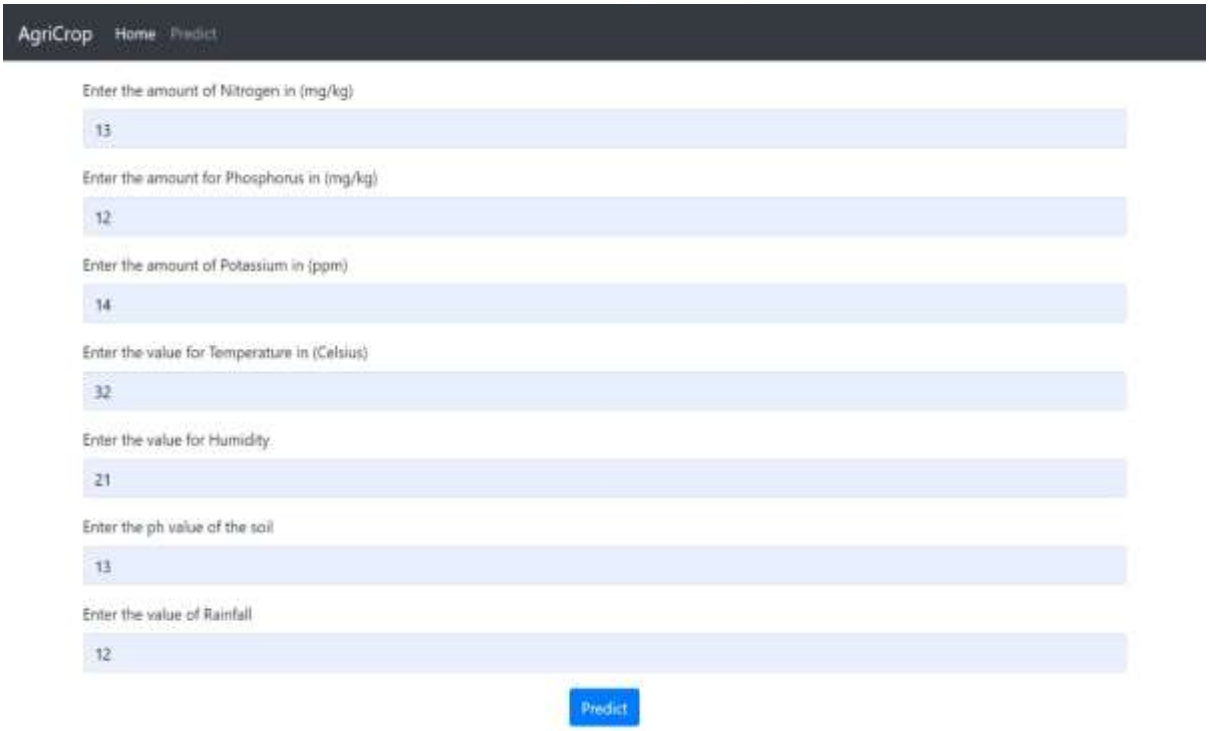

Fig-7: Soil details form

Fig-8: Crop Predicted

6. CONCLUSION:

India is one of many nations whose economy is based primarily on agriculture. Since information technology integration in agriculture will help farmers increase output. The system outlined in this proposed work perates more quickly and provides better prediction accuracy to

determine the best crop for the area. To analyze the crop, it contains a number of soil parameters. This forecast encourages farmers to increase growth and output.

REFERENCES

[1] Kevin Tom Thomas, Varsha S , Merin Mary Saji , Lisha Varghese , Er. Jinu Thomas "Crop Prediction using Machine Learning".

[2] Nischitha K, Dhanush Vishwakarma, Mahendra N, Manjuraju M.R, Ashwini "Crop Prediction using Machine Learning Approaches"

[3] "CROP YIELD PREDICTION USING K-MEANS CLUSTERING" Capstone Design Spring 2020 Amine Bouighoulouden Dr. Ilham Kissani.

[4] "Crop prediction based on soil and environmental characteristics using feature selection techniques" by A. Suruliandi,G. Mariammal & S.P. Raja

[5] "Crop Yield Prediction Using Supervised Machine Learning Algorithm" Hardik Joshi, Monika Gawade, Manasvi Ganu, Prof. Priya Porwal.

[6] Sk Al Zaminur Rahman, S.M. Mohidul Islam, Kaushik Chandra Mitra," Soil Classification using Machine Learning Methods and Crop Suggestion Based on Soil Series",2018 21st International Conference of Computer and Information Technology (ICCIT), 21-23 December, 2018.

[7] S. Panchamurthi. M.E.,M.D. Perarulalan,A. Syed Hameeduddin,P. Yuvaraj,"Soil Analysis and Prediction