International Journal of Research in Engineering, Science and Management
Volume 5, Issue 3, March 2022
https://www.ijresm.com | ISSN (Online): 2581-5792

An Efficient Crop Yield Prediction Using Machine Learning

M. A. Manivasagam[1], Polluru Sumalatha[2*], A. Likitha[3], V. Pravallika[4], K. Venkata Satish[5], S. Sreeram[6]
[1]*Head of the Department, Department of Computer Science and Engineering, Siddartha Institute of Science and Technology, Puttur, India*
[2,3,4,5,6]*Student, Department of Computer Science and Engineering, Siddartha Institute of Science and Technology, Puttur, India*

***Abstract*: In normal, agriculture plays a vital role in the Indian economy and agriculture is the spine of India. In India most of the occupation is agriculture. Farmers always try to cultivate the same type of crop using regular methodologies (not technical methods), Farmers don't have awareness about the technical methods to avoid losses in cultivation. They are mostly using chemicals to solve that problem to the crop and taking suggestions from the public. There are more technical methods that are implemented but not used properly to get a high crop yield. But now a day crop yield gets decreased due to environmental effects (temperature, rain) which affect crop production and lead to loss. So, there is a necessity to know the crop yield before cultivation. This project – Will allow us to predict crops based on climatic conditions and soil conditions, which would help us to know the crop yield earlier of harvesting. It would be more helpful for farmers and also in business to take solid decisions before investing the capital. Machine learning is one such advanced technique deployed to predict crop yield in agriculture. Here, the machine learning algorithm is mostly used to predict crop yield to get more accurate predictions in that Random Forest. By analyzing the issues that affect crop yield is that climatic factors such as average temperature, average humidity, average rainfall, and route map for a selected crop for particular features with date specification. This methodology to design and development of crop prediction and crop yield prediction using different machine learning techniques are used such as K-Near Neighbor, Decision Tree, Random Forest Classifier helps to increase yield and subsequent profit of agricultural production. Some of the features like algorithm verification and check, crop prediction using some parameters, crop analysis, and guide for yield. The project aims to provide an easy way to predict the yield by using various environment parameters accurately and efficiently.**

***Keywords*: Region crop, Risk in cultivation stages, Crop-yield prediction, Random Forest, K-Near Neighbor, Decision Tree.**

1. Introduction

Crop production may affect by environmental factors that would be soil ph, the temperature of a particular region. Parameters of the process vary from one to another. The information of the crop that would be region crop, risk factors and more about crop within the agriculture or different fields in environmental sciences.

There are different fields of information are gathered about the crop, different methodologies and techniques are developed by many analyzers in the world. Maximization of state and different researchers about the field of agriculture. This automated design offers continuous information across a range of timescales and can be readily used for greenhouses. Crop yield prediction shows the ability of the random forest technique of machine learning technology to be used for approximation.

Agriculture of information science and different methodologies of machine learning. Using different innovative techniques in agriculture improves technology in the scientific field of the environment. There is a need to improve the field of agriculture throughout the world. There are more technical methods that are implemented but not used properly to get a high crop yield. But now a day crop yield gets decreased due to environmental effects (temperature, rain) which affect crop production and lead to loss. So, there is a necessity to know the crop yield before cultivation. This project Will allow us to predict crops based on climatic conditions and soil conditions, which would help us to know the crop yield earlier of harvesting. It would be more helpful for farmers and also in business to take solid decisions before investing the capital. Machine learning is one such advanced technique deployed to predict crop yield in agriculture. Agriculture moto is that to increase the crop production (crop yield) and quality of crop using healthy fertilizers to support of health. Nature plays a very important part in every life of the living organism. so that's our duty to save and protect our nature healthily. Agriculture is also from nature(farming). In the olden days, farming is the most important and most of the occupation also. Agriculture is the spine of every country's economy due to without agriculture there is no food and there is no expenditure on products so that will lead to a low economy to the country. In India, agriculture is consistently used as an occupation, It predicts the maximum yield of the crops produced at minimum cost. Early detection and management of the problems associated with crop yield indicators can help increase yield and economic growth. In agribusiness, the most inventive and innovative techniques are used to improve business profits. Because of these useful for the organization in undesirable situations to desirable. By influencing regional weather patterns, large-scale meteorological phenomena can have a significant impact on agricultural production. The suicides of farmers are increasing day by day due to a lack of support and losses. In general, the

*Corresponding author: sumagaye67@gmail.com

temperature and rainfall and highly variable which result in uncertain crop yields. Besides its uncertainty, the distribution of rainfall during the crap period is uneven, receiving a high amount of rain, when it is not needed and lacking it when the crop needs it. So, we've got that knowledge of the cultivators and yield of the crop of weather prediction based on climatic parameters in a particular region. The design of This technique helps in the early detection and management of problems associated with crop yield and can increase yield in agricultural production. This technology is less prone to error than other techniques.

It is an effective machine learning tool for predicting crop yields. To get a high harvest using the innovative arrangement in crop production. Due to these unnatural conditions will affect the economy of crop yield. Most of these climatic conditions can avoid losses in agriculture by forecasting the climatic factor then predicting the crop. prediction of crops mainly rice, wheat, corn which play the most important role in national and international programming. Machine learning is an advanced and growing technology approach in throughout the world and it would be most useful in every field to make solid decisions to solve the most difficult problems.

2. Literature Survey

In [1] Shriya Sahu from Computer Science and Application, "An Efficient Analysis of Crop Yield Prediction Using Hadoop Framework Based on Random Forest," (2020). This paper aims to forecast the yield of crops using the Hadoop framework. The problem with this system is, it is complex due to storing and performing with a huge amount of data. By using the Hadoop framework, it is large and different types of data are accessible in the market accurately predict crop yield.

In [2] G. Vishwa1, J. Venkatesh, "Crop Variety Selection Method using Machine Learning" (2019). This paper gives an overview of analyzing crop selection and implementation using different techniques and methods. Machine learning (ML) has emerged Choosing the wrong crop will affect the yield of crop, losses, and different issues. So, there is a need to select crops more efficient manner. This paper is implemented by using the crop variety selection method (CVSM).

In [3] Shikha Nema, "Prediction of Crop Production in India Using Data Mining Techniques" (2018). This paper focuses on the different varieties of parameters like rainfall, mean temperature, area, and area under irrigation on the production of major crops. It aims to predict crop production using data mining techniques. By using different regression algorithms such as Multiple Linear Regression, Regression analysis was used as a predictive modeling technique for the prediction of crop production.

In [4] Kiran Moray, Arunah Pavate, Suyog Nikam's "Crop Yield Prediction Using Random Forest Algorithm" (2021). In the existing system, It will not give certain or high accuracy. This paper using of the cross-validation technique which improves the decisions and gets accurate results in crop yield. The problem in this system is that accuracy is less than compared to other papers and efficiency.

In [5] Ankara Venugopal, Aparna S, Jinsu Mani, Rima Mathew, "Crop Yield Prediction using Machine Learning Algorithms." In previous, farmers were not able to make certain decisions and predictions on which crop to grow. Here, they used various machine learning techniques. This paper provides information about crop yield for the chosen district, which would be helpful to take decisions based on the conditions. The problem in this paper is compatibility and some of the features.

In [6] Y. Jeevan Nagendra Kumar, V. Spandana, V. S. Vaishnavi, K. Neha, V. G. R. R. Devi "Supervised Machine earning Approach for Crop Yield Prediction". Before there is no increase in yield and profits in producers. Here, Farmers get awareness about decision making and cultivating better crops to get better yield. Using decision tree classifiers.

In [7] "Machine Learning convergence for weather-based crop selection" crop demand is gradually increasing due to its effect on the economy. The agriculture field is decreasing randomly due to the effect of decisions. In this weather-based prediction of crops in which there will be a demand in crop yield and also the efficiency of crop formation and control. Using of KNN, Decision Tree for seasonal weather forecasting.

In [8] Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique" Crop selection method (CSM) is used to solve crop selection problem. This paper provides information about the selection of crop which crop can grow in a particular region based on the region parameters, it would select the crop. Prediction of crop done through these conditions. The problem in this system is that it would predict only crop-based on parameters in given data- It will not predict crop yield.

In [9] Solanki Doina Bein's, "Prediction of Crop Cultivation" Climate change and soil quality affect the local and world food production and world economy. This paper provides information about crop cultivation using different regression like K- near neighbor regression. Which would be more useful weather forecasting Pictorial representation (Maps, Graphs).

In [10], "Maximizing the Yielding Rate of Crops using Machine Learning Algorithm." This paper using of different techniques like data mining techniques. It would be more beneficial to take action regarding the field of agriculture, The techniques of regression are Multilinear regression, Naive Bayes, and random forest regressor. This system which going to help farmers to get more yields and better crops. Various algorithms like Naïve Bayes, Data mining techniques.

3. Existing System

Every country's economy would depend on agriculture. Agriculture depends on the population in the country. So, there is the necessity to improve crop prediction accurately to get consistent results in crop yield. The crop selection method (CSM) is used to solve the crop selection problem. It provides the information about the selection of crop which crop can grow in a particular region based on the region parameters, it would select the crop. Prediction of crop done through these conditions. The problem in this system is that it would predict only crop-based on parameters in given data. It will not predict crop yield.

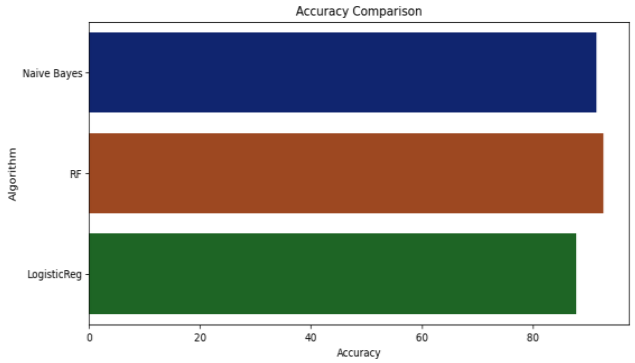

Fig. 1. Accuracy

Table 1
Algorithm's accuracy

Algorithm	Accuracy
Random Forest Classifier	2.81407991690006
Naïve Bayes	91.49621790098573
Logistic Regression	87.82982929223341

The built system is the reliability of prediction in all of the other grains and regions were chosen in the analysis should be 75-87 percent, indicating predictive Accuracy and Performance. In the existing Crop yield, it gives an overview of analyzing crop selection and implementation using different techniques and methods. Machine learning (ML) has emerged Choosing the different issues. So, there is a need to select crops more efficient manner. They used algorithms Naive Bayes, Logistic Regression, random forest algorithm.

They used algorithms Naive Bayes, Logistic Regression, random forest algorithm. There is a need to improve the field of agriculture throughout the world. There are more technical methods that are implemented but not used properly to get a high crop yield. But now a day crop yield gets decreased due to environmental effects (temperature, rain) which affect crop production and lead to loss. So, there is a necessity to know the crop yield before cultivation. This project Will allow us to predict crops based on climatic conditions and soil conditions, which would help us to know the crop yield earlier of harvesting. It would be more helpful for farmers and also in business to take solid decisions before investing the capital.

Uncertainty in agriculture due to climatic change and other diverse parameters, Population out the blast, and sustainable agricultural growth. Decision-making in agriculture plays a huge part in the country. This system Crop Selection Method (CSM) would be the selection of crop, which crop is suited for the particular area based on the conditions. CSM allows a seasonal-oriented Method and area-oriented. The method of crop selection is precision agriculture and precision farming using progressive technologies. Data accumulation through the internet of agriculture precision things of the application interface, these methods were then re-tuned with the additional features and compared to the models.

Different techniques would be useful for accumulating features in crop selection. Here 80% of the dataset is taken as a training dataset and 20% for testing.

4. Proposed System

The main aim of this system (Crop Yield Prediction) is to find the dataset which contains production details and also details about soil parameters and also environment parameters like rainfall, temperature, Humidity, Ph, etc. Thus, various factors are assessed and the factors strongly lead to accurate prediction of the production of the crop. And then using predicted crop and land area in a hectare will predict crop yield. using various Machine Learning algorithms like Decision trees, KNN, Random Forest classifiers. By using Random Forest Classifier algorithm, we might be Increase accuracy from (87%) to 95% and explore features like algorithm verification and check, crop prediction using some parameters, crop analysis, and guide for yield. High Efficiency, Extra features like the season, season duration, Region crop, Risk in cultivation, Minimum water for cultivation, Maximum water for cultivation, Minimum time for cultivation, Maximum time for cultivation, and Recommended pesticides for the crop.

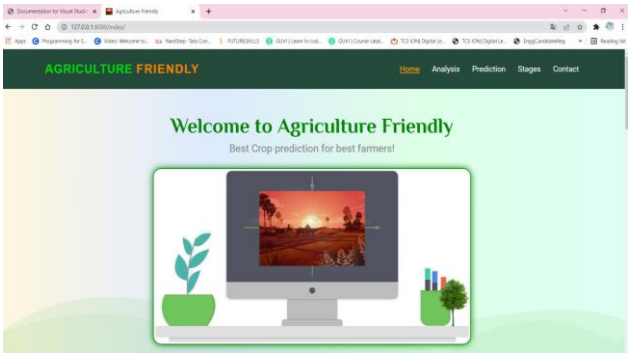

Fig. 2. Website for crop yield prediction

	Temperature	Humidity	Ph	Rainfall	Label
2	20.87974371	82.00274	6.502985	202.9355	rice
3	21.77046169	80.31964	7.038096	226.6555	coffee
4	23.00445915	82.32076	7.840207	263.9642	sugarcane
5	26.49109635	80.15836	6.980401	242.864	millet
6	20.13017482	81.60487	7.628473	262.7173	peas
7	23.05804872	83.37012	7.073454	251.055	jute
8	22.70883798	82.63941	5.700806	271.3249	cotton
9	20.27774362	82.89409	5.718627	241.9742	groundnut
10	24.51588066	83.53522	6.685346	230.4462	maize
11	23.22397386	83.03323	6.336254	221.2092	watermelon
12	26.52723513	81.41754	5.386168	264.6149	lentil
13	23.97898217	81.45062	7.502834	250.0832	grapes
14	26.80079604	80.88685	5.108682	284.4365	apple
15	24.01497622	82.05687	6.984354	185.2773	banana
16	25.66585205	80.66385	6.94802	209.587	pigeon peas
17	24.28209415	80.30026	7.042299	231.0863	black grams
18	21.58711777	82.78837	6.249051	276.6552	mango
19	23.79391957	80.41818	6.97086	206.2612	kidney beans
20	21.8652524	80.1923	5.953933	224.555	tobacco
21	23.57943626	83.5876	5.853932	291.2987	wheat
22	21.32504158	80.47476	6.442475	185.4975	Mung Bean
23	25.15745531	83.11713	5.070176	231.3843	papaya
24	21.94766735	80.97384	6.012633	213.3561	pomegrante
25	21.0525355	82.6784	6.254028	233.1076	coconut
26	23.48381344	81.33265	7.375483	224.0581	chick peas
27	25.0756354	80.52389	7.778915	257.0039	rice
28	26.35927159	84.04404	6.2865	271.3586	rice
29	24.52922681	80.54499	7.07096	260.2634	rice

Fig. 3. Dataset

Random Forest Algorithm is an algorithm for machine learning which belongs to the technique of supervised learning. It can be used for problems with Regression in Machine Learning mostly used for crop prediction in this system. Machine learning makes models easier and more accurate from the complex natural system with many inputs. It can be used to estimate crop prediction in a long and short time. The previous data is also important that that information only given to the training the data. So, there is a need to collect accurate information from standard sources. The data should be precise and efficient of agriculture parameters. To perform the prediction of crop yield more accurately and efficiently. prediction of crop yield will be trained with gathered data (accurate) and consistent data, data in average temperature, soil quality, soil Ph, and rainfall various machine learning classifiers. By applying the above machine learning algorithms, the Random Forest algorithm provides more accurate results than compared to other algorithms. The system is trained by the gathered data of the model of the random forest then it will predict the crop based on parameters then it predicts crop yield based on crop prediction and area predicts crop prediction from the gathering of previous data. Using previous information on weather, soil, temperature, and several other factors the data is given. The website which we developed runs the algorithm on a random forest classifier and shows the list of crops and stages of the crop as a route map. Based on suitable for entered data with predicted yield.

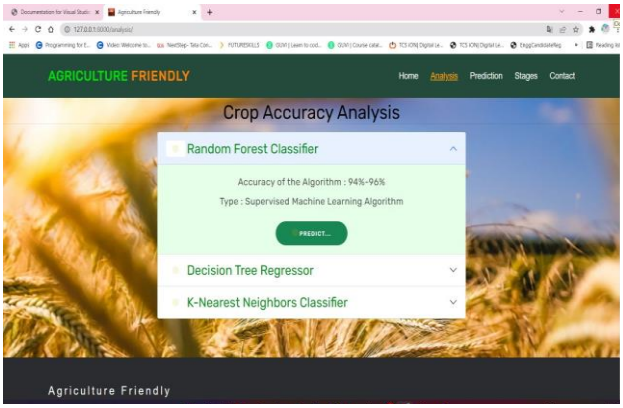

Fig. 4. Algorithm analysis

The Initial step of the process is that collect data of agriculture based on some parameters and then preprocess the data which would clean the noisy and null data in gathered data to form as preprocessed data and then train the data using various models of machine learning then predict the crop on weather and climatic conditions to get crop productions. Here, the model used for this system is the random forest which gets accurate results. Parameters like Temperature, Humidity, soil Ph, rainfall then it will predict crop which suits our land-based on those parameters. After prediction of the crop, now parameters to predict crop yield is that, predicted crop and are in a hectare. Then it will predict crop yield for our area. Implement route map (stages of the crop with Date specification) for the selected crop in a particular season. Pictorial representation of crop yield. Here 70% of the dataset is taken as a training dataset and 30% for testing.

5. System Architecture

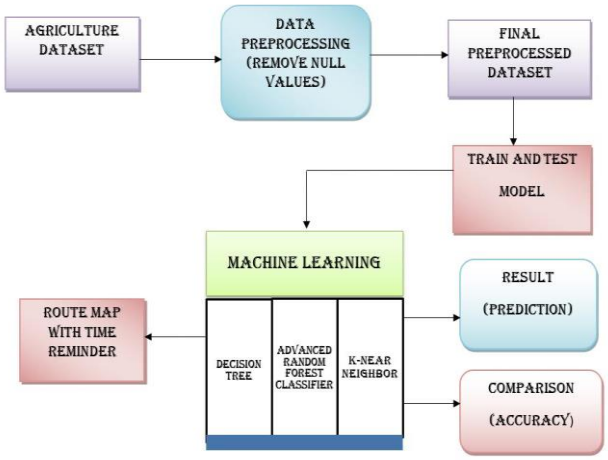

Fig. 5. Block diagram

The above block diagram explains how the system interacts with the elements in the architecture. The initial step is that collect an agriculture dataset with some parameters of soil and climatic conditions. Then preprocess the data which would clean the noisy and null data in gathered data to form as preprocessed data that would be final preprocessed data, then it would train and test the final preprocessed data and then train the data using various models of machine learning like Random Forest classifier, KNN, Decision tree. By using this technique first train the data and then predict the crop. Here, Random Forest algorithms give the finest set of results (95%). During preprocessing step split the dataset into training (70%) and testing (30%) datasets. Using supervised machine learning in this system to predict crops and then predict crop yield. Then comparing the accuracy of three models of supervised machine learning. Decision tree (90%), KNN (85%), and the random forest get more accuracy (95%), So, users can use algorithms based on their wish, It's optional. Results depend on the algorithm. Then route map with date specification, it's like crop stages from the initial step to final step of crop cultivation. Some extra features in the system are that predict that crop and details regarding that predicted crop. season, season duration, Region crop, Risk in cultivation, Minimum water for cultivation, Maximum water for cultivation, Minimum time for cultivation, Maximum time for cultivation, and Recommended pesticides for the crop.

6. Results

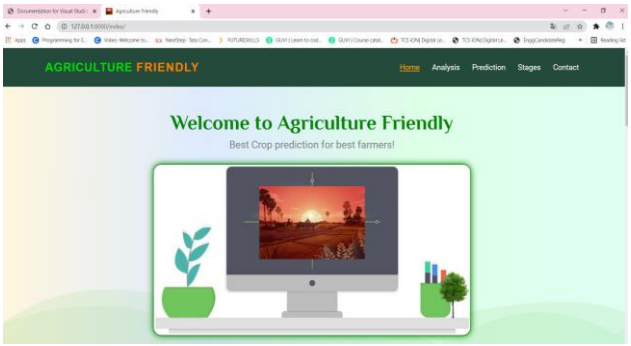

Fig. 6. Agriculture friendly

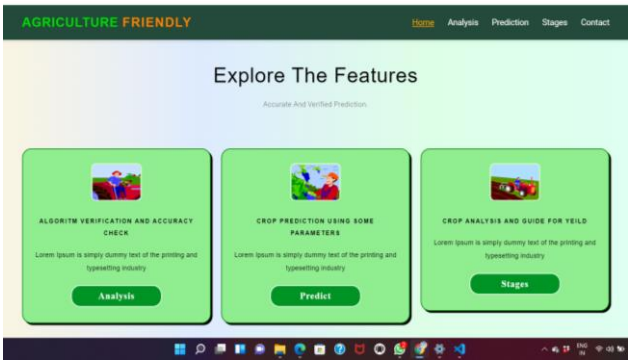

Fig. 7. Features

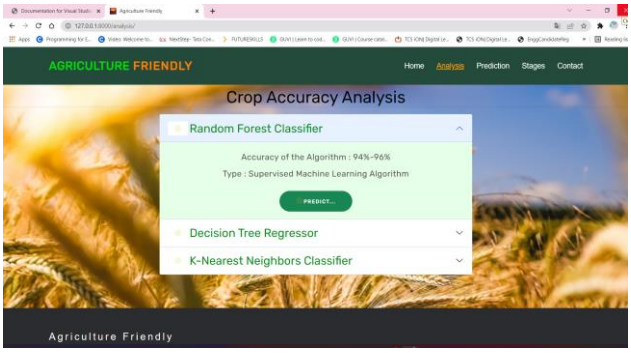

Fig. 8. Crop accuracy analysis

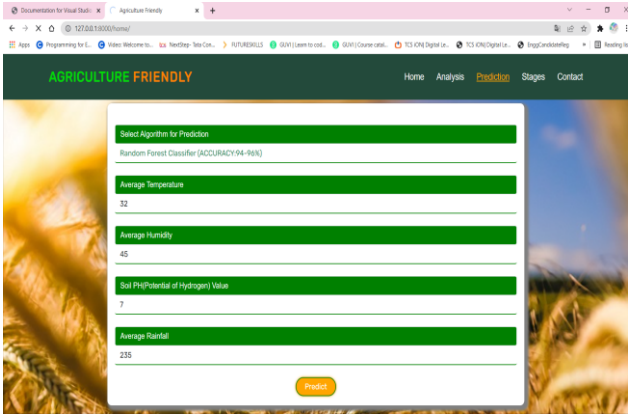

Fig. 9. Crop prediction

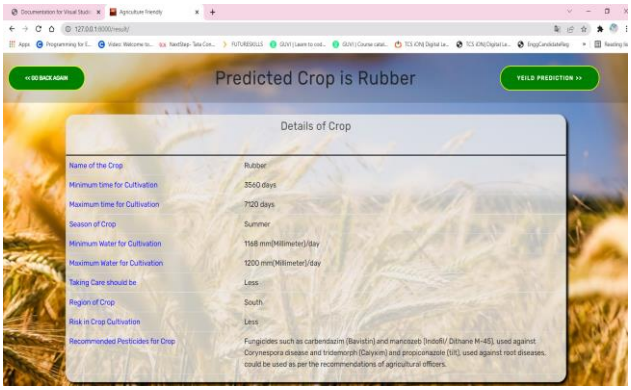

Fig. 10. Predicted crop details

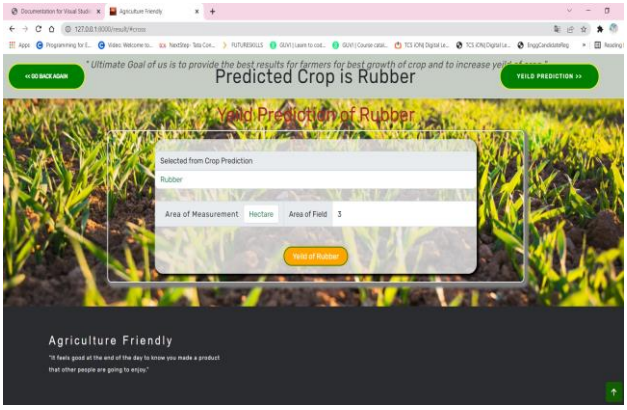

Fig. 11. Crop yield

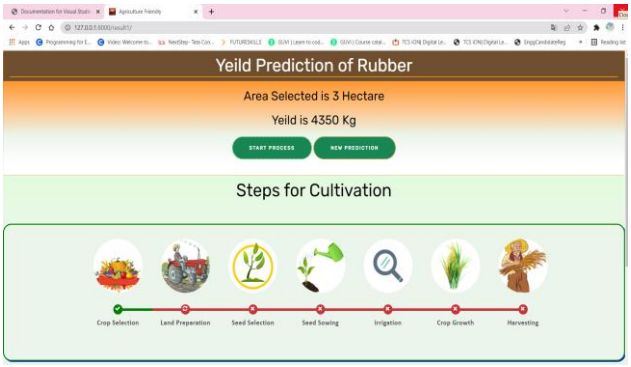

Fig. 12. Prediction of crop yield

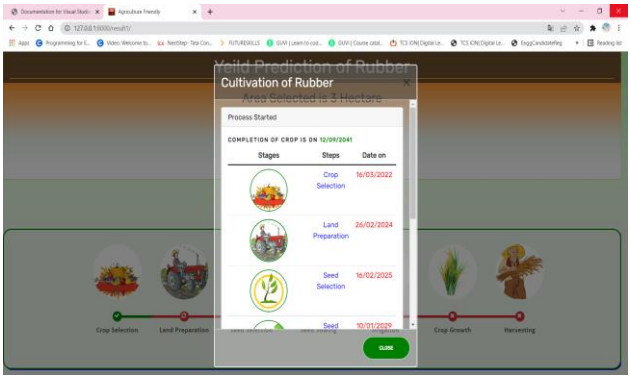

Fig. 13. Crop cultivation

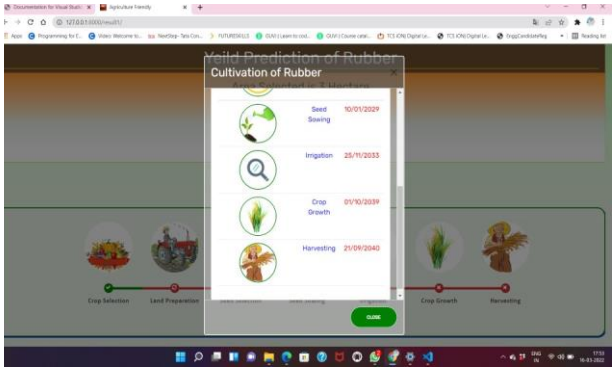

Fig. 14. Steps with date specification

7. Conclusion

This system will predict crop and crop yield, based on the analysis of crop details i.e., gathered data of agriculture. This paper can conclude that the use and implement the system based on machine learning techniques to get more accurate results to make decisions for the farmers especially. Decision-making plays a critical role in agriculture to get profit or losses, so there is more important to take the consistent decision. Using a random forest algorithm makes to predict crops consistently and efficiently with a lack of errors in that model. By analyzing the issues that affect crop yield is that climatic factors such as average temperature, average humidity, average rainfall, and route map for a selected crop for particular features with date specification. This methodology to design and development of crop prediction and crop yield prediction using different machine learning techniques are used such as K-Near Neighbor, Decision Tree, Random Forest Classifier helps to increase yield and subsequent profit of agricultural production. Some of the features like algorithm verification and check, crop prediction using some parameters, crop analysis, and guide for yield. The project aims to provide an easy way to predict the yield by using various environment parameters accurately and efficiently.

8. Future Enhancement

In the coming years, can try applying market trends. Using different innovative techniques in agriculture improves technology in the scientific field of the environment. There is a need to improve the field of agriculture throughout the world. and also, we can add some more features like irrigation, water levels in the area as parameters, different graphical representations of irrigation level, and a performance measurement graph too. Risk management precautions to the user.

References

[1] Shriya Sahu, “Analysis of Crop Yield Prediction Using Hadoop Framework Based on Random Forest,” 2020.
[2] G. Vishwa1, J. Venkatesh, C. Geetha, “Crop Variety Selection Method using Machine Learning,” 2019.
[3] Shikha Nema, “Prediction of Crop Production in India Using Data Mining Techniques,” 2018.
[4] Kiran Moray, Aruna Pavate, Suyog Nikam, “Crop Yield Prediction Using Random Forest Algorithm,” 2021.
[5] Anakha Venugopal, Aparna S, Jinsu Mani, Rima Mathew, “Crop Yield Prediction using Machine Learning Algorithms,” 2021.
[6] Y. Jeevan Nagendra Kumar, V. Spandana, V. S. Vaishnavi K. Neha, V. G. R. R. Devi, “Supervised Machine earning Approach for Crop Yield Prediction,” 2020.
[7] Sonal Jain, “Machine Learning convergence for weather-based crop selection,” 2021.
[8] Rakesh Kumar, M. P. Singh, Prabhat Kumar, and J. P. Singh, “Crop Selection Method to Maximize Crop Yield Rate using Machine Learning Technique,” 2018.
[9] Raxitkumar Solanki Doina Bein, “Prediction of Crop Cultivation,” 2019.
[10] Yogita Masare, Sneha Mahale, Manjusha Kele, Ashvin Upadhyay, Bhushan R. Nanwalkar, “The System for Maximizing the Yielding Rate of Crops using Machine Learning Algorithm,” 2021.