



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume: 11    Issue: III    Month of publication: March 2023**

**DOI: <https://doi.org/10.22214/ijraset.2023.49877>**

**[www.ijraset.com](http://www.ijraset.com)**

**Call:  08813907089**

**E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)**

# Prognosis of Crop Yield Using Machine Learning Techniques

Veera Jayanth Kuppam<sup>1</sup>, Arisetty Sai Badrinath<sup>2</sup>, Yellaturi Vijay<sup>3</sup>, G. Haranadha Babu<sup>4</sup>

<sup>1, 2, 3</sup>Undergraduate Students, Department of Information Technology, KKR & KSR Institute of Technology and Sciences, India

<sup>4</sup>Asst. Professor, Department of Information Technology, KKR & KSR Institute of Technology and Sciences, India

**Abstract:** Agriculture is the major occupation of many people in India. India was standing in the leading producer of agricultural goods. As the population of India increases gradually, the yield of the crop may be not sufficient in further days. Also, the Indian economy is based on agriculture. Using the technologies like Machine Learning, Deep Learning, and Artificial Intelligence we can prognosis the yield of the crop using some parameters like rainfall, temperature, crop to be cultivated, and pesticides. This prediction helps the farmers to know the yield before they cultivate.

**Keywords:** Prognosis, Crop Yield, Agriculture, Random Forest, Machine Learning.

## I. INTRODUCTION

Agriculture was the first occupation of human beings which was developed before 10,000 years approximately. It was the one that helped to start our human civilization. As per the statistics up to 2018, nearly 50% of the population is using agriculture as their occupation.

India was the second leading producer of agricultural goods. Even in 2020, the majority of the population in our country is dependent on Farming. Also, agriculture has contributed nearly 17 to 18% of the Gross Domestic Product (GDP) of India. In India majority of the land space was used for Agriculture, which pushed all other countries and took the top place in the list. Also, India was exporting these Agricultural Goods to other countries which were nearly \$3.5 billion (in March 2020). India was the seventh largest exporter of Agricultural Goods for more than 120 countries namely Japan, Southeast Asia, SAARC Countries &, etc....

As of now, the Agriculture products are sufficient enough for the population present. But as per the statistics, the population may increase by nearly 25% which causes a crisis in Agricultural Goods. Also, the weather conditions are getting unpredictable due to the pollution, global warming, etc.. which ruins the crop sometimes, leads to loss of the yield. So the crisis of the crop increases more. due to these reasons even if one yield misses it will be a big loss. Due to this, we must use the latest technologies to increase the yield of the crop which helps future generations.

Machine learning was the technology that takes past data as input & trains a model accordingly, when the user gives the new data the system analyzes the data as per the model it created during the training process & predicts the required output. Using this Technology we can implement a system to forecast or prognosis the yield of a crop which helps to increase the production & also the economy. There are many existing systems which as using the techniques like classification & Regressions, Support Vector Machines, Decision trees, etc..., whose accuracy was low as compared with Random Forest Algorithm In this Algorithm the data given to the system was divided & each data is transformed into the decision tree. Later the output is given by a voting process which further increases the accuracy. So, this helps to predict the yield with high accuracy which helps farmers economically and also.

## II. LITERARY SURVEY

Machine Learning can be used for predicting the crop yield as the tool for taking decisions using some parameters. It can be used as the tool to increase the crop yield. There are no. of techniques in many domains to support the yield prediction research. According to the research, we are using the Kalman filter Algorithm. to extract the features using linear discriminant Analysis. [1]

Yield Prediction can be done using ML techniques using the parameters like yield, rainfall & geographical Conditions. ML techniques like Random Forest Algorithm can be used for both classification and regression. Weather and other condition affect the crop yield, which yields less. Machine learning nowadays is used worldwide as it is successful in many fields. Using this research, Farmers can estimate the yield of the crop. [2]

The population of India was increasing day by day which increases the demand for crops which directly relates to agriculture. There are many advanced Technologies like remote sensing, image processing, and Data mining. etc. Parameter like average rainfall, average temperature, etc... plays an important role while determining the yield of the crop.[3]

The search aims to determine the yield of the Soyabean using the Synergy of satellite, ML & Unmanned Aerial Vehicles (UAVs). The data needed is collected up to 2017 from various countries. Canopy spectral features & UAV images were combined to predict the Soyabean yield. Parameters like canopy height & canopy cover were used to predict the yield of soybean.[4].

The research on crop yield prediction mainly aims the palm oil. They are various techniques of machine learning, that are used for predicting the yield of the product. Firstly we need to collect the data from the palm oil plant & latex using machine learning to make a model. We use techniques like DNN, LSM & etc. [5].

This project helps the farmers to predict the yield of the crop. This includes the domains like Machine Learning, Web Technology, etc. We made a page where the farmer can enter the values of parameters to know the prediction of yield. In India, as the population is increasing the demand for crops increases. Using this tool we can predict the yield in advance. [6]

The increase in population directly relates to the demand for agricultural products to reach the hunger of everyone. Technology can help the former to increase the yield of the crop which helps them reach the demand for agricultural products. We used the Decision tree algorithm of machine learning which helps & makes it easy to predict the yield. This is very beneficial for small foramen, [7]

The desired outputs of the given inputs will be given only when the data related to it is collected & combined in the best way. Parameters like the ph of the soil, rainfall, and product name can be used to get the desired output as the Random Forest algorithm is used to function as it to the one techniques having high we can also predict the type of the crop which yields the highest. [8].

The project also includes parameters like the State, district, season, area and to predict the crop. In this, we added all the types of products (like crops) cultivated in India. Keeping these we used the algorithm named stacked regression to predict the most advanced or faster way. We added features like regional languages that help all the people using it. [9].

The parameters like water, UV rays, pesticides and others are used to predict the yield. To predict we used two algorithms like Support Vector Regression & Linear Regression functions. The functions are used separately to find out the error rate. The best function is chosen for prediction. [10]. Agriculture can be represented as the backbone of India as it is performed by many of the people and also it decides the growth of the Indian economy, which shows a drastic change in Domestic Production. So, we are implementing a project to predict the yield of the crop in advance. In the future, it can be also used for suggesting fertilizers, arable land and crops. In addition, the sunlight sources and crop health is monitored regularly to achieve better yields.[11].

The paper predicts yields of about all types of crops grown in India. The script is novel with simple parameters such as state, district, season, region, etc., allowing the user to predict the harvest for the desired year. The output of the yield is provided using the web page in the browser, but in future the entire project can be implemented using the mobile app were all the users can access it using their preferred local or regional language.[12]

The climate changes in India effects the yield of the crop, which directly effects the storages the the crop. After analysing the data of 20 years, we prepared a function. This allows the farmer or the user to predict the yield in advance which helps him in marketing and to care of storage measures. We design a web application to predict the yield in a user friendly way.[13]

Agriculture is the occupation which is not only for the economic purpose also it deals with the survival of the human being on the earth. Crop Yield Prediction is not an easy task; it includes various parameters like sunlight, water, humidity, etc. Uncontrollable environmental parameters are helpful in understanding the impact on yield, and further evaluation in a controllable environment will suffice to meet the routine need for yield measurements.[14].

### III. METHODOLOGY

#### A. Data Collection

As, the Machine Learning algorithm needs the old data to prepare a model for user data input we must collect the data from various sources. For this research we have collected data from Kaggle website i.e. yield, rainfall, temperature, pesticides data, etc... that are needed for our research.

#### B. Data Preprocessing

The improper format or the empty values of the data disturb the model which will effect the prediction of the user input. So, the data must be preprocessed before it is sent or used by the Machine Learning Algorithm. The Data Preprocessing has many processes like, Data Cleaning, Data Transformation, Data Reduction. This process is the conversion of the raw data into the useful and efficient format.



### C. Machine Learning

The main phase of the research is Machine Learning. The old collected and processed data is given as input to the Algorithm and the prediction model is given as output. There are many algorithms in this domain such as Linear regression, Logistic regression, Random Forest, Decision tree, Support Vector Machine algorithm, Naive Bayes algorithm, etc...

Random Forest Algorithm is the algorithm which is having the high accuracy. This process involves the following :

- 1) Select the random samples for the given training data.
- 2) Constructing the Decision Tree for each training data sample.
- 3) Voting is done on averaging the decision tree output.
- 4) The output with high voting is given as output.

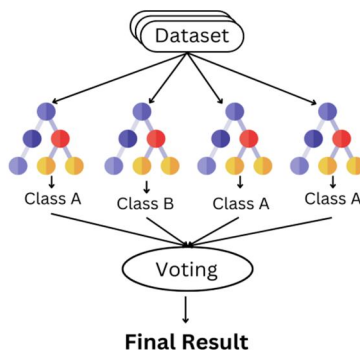
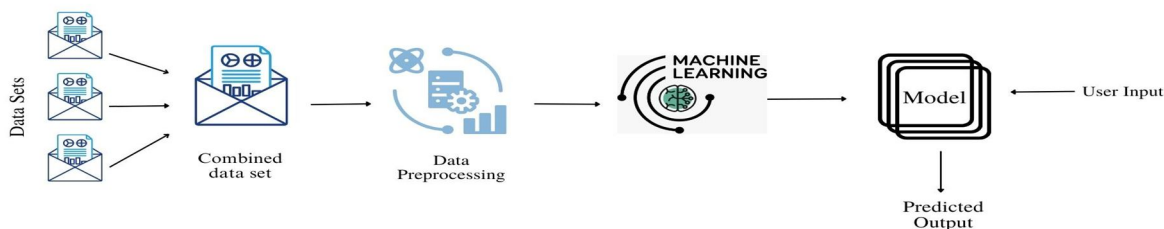


Fig 3.1 Random Forest Algorithm

For all this to be used, we are implementing a web application through which a user (like farmer etc..) can use our application. We used a web framework called Flask i.e. Python Backend Framework which depends mainly on the handlers to navigate and perform operations.

For this purpose we have to enable a new environment and install the flask packages.

## IV. SYSTEM ARCHITECTURE



- 1) *Data Sets*: Data set is a collection of various data that is used to train the model. In our project we used datasets like rainfall, temperature, soil data, yield data of previous years.
- 2) *Combined Dataset*: It is a combination of all data into a single document which is used to train the model.
- 3) *Data Preprocessing*: Data preprocessing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model. When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way.
- 4) *Machine Learning*: In this phase the data is given to the system to create a model for predicting the result for newer data given by the user.
- 5) *Model*: It is created in the earlier phase, this is the used for the prediction purpose of the newer data given as the input by the user.

### V. RESULTS

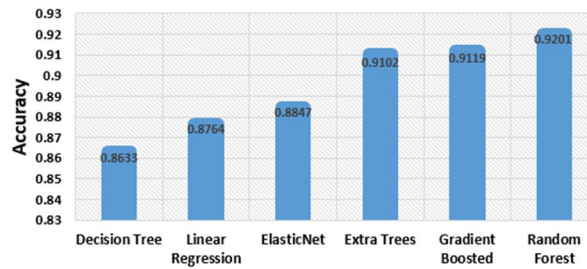


Fig 5.1 Accuracy of Machine Learning Techniques

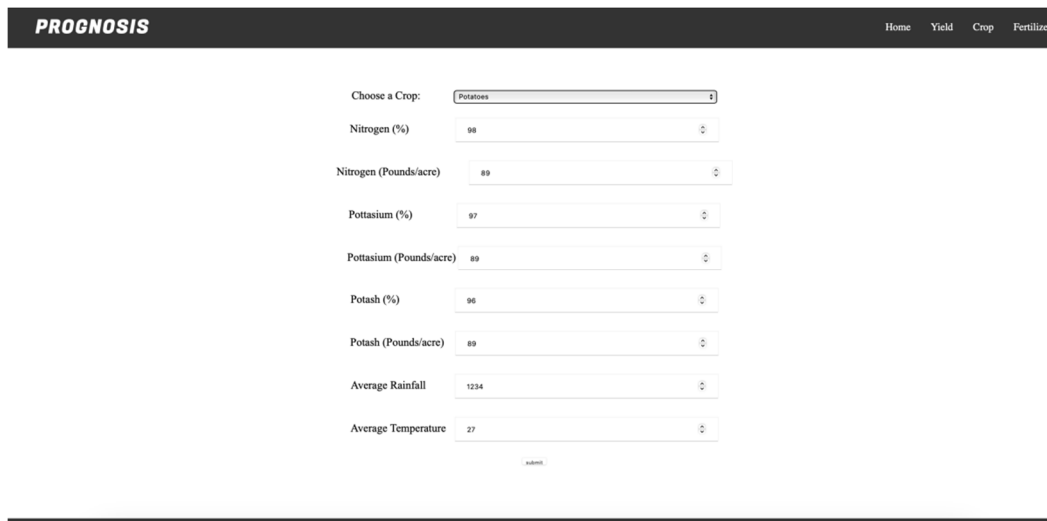


Fig5.2 Web Page Design

When the user enters the value in the fields of the above form, the model which is already prepared will be called along with the values and the yield of the crop will be predicted in hg/ha and displayed in the another page.

### VI. CONCLUSIONS

In our research, we used Random Forest Algorithm which is having high accuracy and speed further, we can add the parameters like soil data, and can also use the drone for capturing the data in real time. We can also develop a mobile application in all regional languages and help the farmers with all the features like yield prediction, crop recommendation, pesticide recommendation, disease identification, etc.

### REFERENCES

- [1] Swati Vashisht, Praveen Kumar, Praveen Kumar, Improved Extreme Learning Machine for Crop Yield Prediction, 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)
- [2] Ranjani J, V.K.G Kalaiselvi, A. Sheela, Deepika Sree D, Janaki G, Crop Yield Prediction Using Machine Learning Algorithm, 2021 4th International Conference on Computing and Communications Technologies (ICCT),
- [3] Sunil G L, Nagaveni V, Shruthi U, A Review on Prediction of Crop Yield using Machine Learning Techniques, DOI:10.1109/TENSYMP54529.2022.9864482, ISBN:978-1-6654-6659-2
- [4] Maitiniyazi Maimaitijiang, Vasit Sagan, Felix B. Fritschi, Crop Yield Prediction using Satellite/Uav Synergy and Machine Learning, 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, DOI: 10.1109/IGARSS47720.2021.9554735, Print on Demand(PoD) ISBN:978-1-6654-4762-1
- [5] Mamunur Rashid, Bifta Sama Bari, Yusri Yusup, Mohamad Anuar Kamaruddin, Nuzhat Khan, A Comprehensive Review of Crop Yield Prediction Using Machine Learning Approaches With Special Emphasis on Palm Oil Yield Prediction, IEEE Access ( Volume: 9), DOI: 10.1109/ACCESS.2021.3075159
- [6] Namgiri Suresh, N.V.K. Ramesh, Syed Inthiyaz, Crop Yield Prediction Using Random Forest Algorithm, 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), DOI: 10.1109/ICACCS51430.2021.9441871
- [7] Ms Kavita, Pratistha Mathur, Crop Yield Estimation in India Using Machine Learning, 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), DOI: 10.1109/ICCCA49541.2020.9250915



- [8] Y. Jeevan Nagendra Kumar, V. Spandana, V.S. Vaishnavi, K. Neha, V.G.R.R. Devi, Supervised Machine learning Approach for Crop Yield Prediction in Agriculture Sector, 2020 5th International Conference on Communication and Electronics Systems (ICCES), DOI: 10.1109/ICCES48766.2020.9137868
- [9] Potnuru Sai Nishant, Pinapa Sai Venkat, Bollu Lakshmi Avinash, B. Jabber, Crop Yield Prediction based on Indian Agriculture using Machine Learning, 2020 International Conference for Emerging Technology (INCET), DOI: 10.1109/INCET49848.2020.9154036, ISBN:978-1-7281-6222-5
- [10] Fatin Farhan Haque, Ahmed Abdelgawad, Venkata Prasanth Yanambaka, Kumar Yelamarthi, Crop Yield Analysis Using Machine Learning Algorithms, 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), DOI: 10.1109/WF-IoT48130.2020.9221459, ISBN:978-1-7281-5504-3
- [11] M. Kalimuthu, P. Vaishnavi and M. Kishore, "Crop Prediction using Machine Learning," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020, pp. 926-932, doi: 10.1109/ICSSIT48917.2020.9214190.
- [12] P. S. Nishant, P. Sai Venkat, B. L. Avinash and B. Jabber, "Crop Yield Prediction based on Indian Agriculture using Machine Learning," 2020 International Conference for Emerging Technology (INCET), 2020, pp. 1-4, doi: 10.1109/INCET49848.2020.9154036.
- [13] N. Suresh et al., "Crop Yield Prediction Using Random Forest Algorithm," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021, pp. 279-282, doi: 10.1109/ICACCS51430.2021.9441871.
- [14] F. F. Haque, A. Abdelgawad, V. P. Yanambaka and K. Yelamarthi, "Crop Yield Analysis Using Machine Learning Algorithms," 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), 2020, pp. 1-2, doi: 10.1109/WF-IoT48130.2020.9221459.



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)