IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, August 2022
Impact Factor: 6.252

# Crop Yield Prediction of Indian Agriculture using Machine Learning Algorithms

**Ms. Bhoomika P**
Student, Department of Computer Science and Engineering
REVA University, Bangalore, Karnataka, India

**Abstract:** *Today approximately 40% of the living area in the world is conquered by crop agriculture. Agriculture is the backbone of our economy. The practice of producing plants and livestock is known as agriculture. The history of agriculture began a several thousand years ago and is still being practiced at a wide range. Machine Learning is one of the major tool which supports crop yield prediction which may include to decide on what crop to be grown during a particular season of the crops. We can predict the yield of the crop using the machine learning concept Classifiers such as KNN (Kernel - Nearest Neighbour), SVM (Support Vector Machine), Logistic Regression, Naïve Baye's Theorem, Random Forest Classifier. This paper predicts the crop yield using KNN Algorithm. According to the analysis, the most considered factors in our prediction are : Production of the crop, Area of the crop yield produced. The prediction made by machine learning algorithms will help the farmers to come to a decision which crop to grow in their agricultural land to induce the most yield. This helps to bridge the gap between technology and agriculture sector. Consumers would suffer greatly if agriculture were to cease, as their diets would become unbalanced. Food imports from other nations would be necessary as our nation would have to rely on them for its food supply.*

**Keywords:** Agriculture, Analysis, Crop yield prediction, KNN Algorithm, Machine Learning, Agriculture Sector.

## I. INTRODUCTION

The Indian economy is based on agriculture. Agricultural yield in India is mostly determined by meteorological conditions. Rice cultivation is primarily reliant on rain. Farmers will benefit from timely guidance and analysis to estimate future crop output and to assist them maximise crop yield. Predicting crop yields is a significant agricultural issue. Farmers used to forecast their production based on previous year's yield results. Thus, different methodologies or algorithms exist for this type of data analytics in crop prediction, and we can predict crop yield using such algorithms.
The algorithm utilised is the random forest algorithm. There is a growing range of applications and the function of Big data analytics techniques in agriculture using all of these algorithms and the inter-relationship between them. The agricultural area has been slowly declining since the invention of new creative technology and procedures.
Seasonal climatic conditions are also being altered as a result of these cultivating techniques, posing a threat to key assets such as land, water, and air, resulting in food insecurity. By analysing all of these challenges and problems, such as weather, temperature, and a variety of other elements, we have discovered that there is no appropriate answer or technology to help us overcome the scenario we are in. In India, there are numerous options for increasing agricultural economic growth. There are numerous methods for increasing and improving agricultural output and quality.
Crop yield production can also be predicted via data mining. The major goals are :

1. To predict agricultural yield using machine learning techniques.
2. To provide a user interface that is simple to use.
3. To improve crop yield forecasting accuracy.
4. To investigate various climate variables (cloud cover, rainfall, temperature).

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, August 2022
Impact Factor: 6.252

## II. LITERATURE SURVEY

[1] This paper will assist the farmers in learning the yield of their crop, enabling them to make the best choices before cultivating on the agricultural field.By creating a working prototype of an interactive prediction system, it tries to find a solution.

[2] This paper proposesadvanced regression techniques including Kernel Ridge, Lasso, and ENet algorithms and leverages the idea of stacking regression to improve the algorithms and provide a more accurate prediction.

[3] The major goal of this research is to use several machine learning approaches to forecast the agricultural production. By taking into account variables like temperature, rainfall, area, and other characteristics, the predictions provided by machine learning algorithms will assist farmers in choosing which crop to cultivate to induce the greatest yield.

## III. METHODOLOGY

Data is an essential component of every Machine Learning system. To put the system in place, we chose the Indian state of Maharashtra. It was required to collect data at the district level because the climate varies from place to place. To put the system in place, historical data on the crop and the climate of a certain location was required. This information was gleaned from a variety of government websites. www.data.gov.in was used to acquire information on Maharashtra's crops, and www.imd.gov.in was used to obtain information on the state's climate. Precipitation, temperature, cloud cover, vapour pressure, and wet day frequency are the climatic variables that have the greatest impact on the crop. As a result, data on these climatic indicators was collected on a regular basis.

- **Dataset Collection:** We collect data from numerous sources and prepare datasets during this phase. In addition, the provided dataset is being used for analytics (descriptive and diagnostic). Data.gov.in and indiastat.org are two websites that provide online abstracts. The yearly abstracts of a crop will be used for at least ten years. Anarchic time series behaviour is often accepted in these datasets. The primary and necessary abstracts were combined. Random Forests for Crop Yield Predictions at the Global and Regional Level.
- **Data Partitioning:** The entire dataset is divided into two sections: for example, 75% of the data is used to train the model, while the remaining 25% is utilised to test the model.
- To make predictions about the future Algorithms for Machine Learning:
- **Supervised learning:** Using labelled examples, supervised machine learning algorithms may apply what has been learned in the past to fresh data. The system may provide targets for any new input after sufficient training. The learning algorithm can also discriminate its outputs from the correct, intended output and detect mistakes in order to alter the model accordingly.
- **Unsupervised machine learning** techniques, on the other hand, are utilised when the material required to train is neither tagged nor categorised. Unsupervised learning investigates how computers might infer a function from unlabelled data to describe a hidden structure. The system does not figure out the proper output to explain hidden structures from unlabelled data; instead, it studies the data and can make inferences from datasets.
- **Random Forest Classifier:** Random forest is the most popular and powerful supervised machine learning algorithm capable of performing both classification and regression tasks. It works by training a large number of decision trees and then generating outputs of the class that is the mode of the classes (classification) or mean prediction (regression). The more trees there are in a forest, the more accurate the prediction becomes.

## IV. RESULT

This paper uses machine learning approaches to improve crop output. The method that produces high accuracy forecasted the yield of the correct crop. With input libraries like Scikit-Learn, Numpy, Matplotlib, and Pandas, Python 3.8.5 (Jupyter Notebook) is used to implement the machine learning algorithms. An Android application that was created questioned the outcomes of the machine learning analysis. The crop name and accompanying yield were displayed through the Jupyter Notebook itself.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, August 2022
Impact Factor: 6.252

## V. FUTURE WORK

We might try using a data independent system in the upcoming years. No matter the format, our system must function with the same accuracy. Integrating soil information into the system is advantageous since crop selection also considers soil knowledge as a factor. Crop cultivation also requires effective irrigation. Rainfall might indicate whether or not additional water availability is required. By making this study available to all of India, it may be improved to a higher degree.

## VI. CONCLUSION

The current study demonstrated the possible application of data mining approaches in predicting agricultural output based on meteorological input characteristics. The prediction accuracy is greater than 75% in all of the crops and districts studied, demonstrating higher forecast accuracy. The user-friendly approach for estimating crop yield can be utilised by any user with their choice of crop by giving climate data for that location.

## REFERENCES

**[1].** MayankChampaneri, DarpanChachpara, ChaitanyaChandvidkar, MansingRathod,"Crop Yield Prediction Using Machine Learning", International Journal of Science and Research.

**[2].** PotnuruSaiNishant, PinapaSaiVenkat, Bollu Lakshmi Avinash, B. Jabber, "Crop Yield Prediction based on Indian Agriculture using Machine Learning", International Conference for Emerging Technology (INCET).

**[3].** AnakhaVenugopal, Aparna S, Jinsu Mani, Rima Mathew, Vinu Williams, "Crop Yield Prediction Using Machine Learning Algorithms", International Journal of Engineering Research and Technology (IJERT).