

e-ISSN: 2582-5208

#### International Research Journal of Modernization in Engineering Technology and Science (Peer-Reviewed, Open Access, Fully Refereed International Journal)

Volume:05/Issue:03/March-2023

Impact Factor- 7.868

www.irjmets.com

# CROP YIELD PREDICTION USING DECISION TREE, RANDOM FOREST AND GRADIENT BOOSTING REGRESSION TECHNIQUES

B. Harika<sup>\*1</sup>, N. Gayathri<sup>\*2</sup>, S. Prasanthi<sup>\*3</sup>, G. Sai Pramod<sup>\*4</sup>, K.Purna Teja<sup>\*5</sup>

\*1,2,3,4,5 Dept. Of Compute Science And Engineering, Aditya Institute Of Technology And

Management, Tekkali, India.

### ABSTRACT

Machine learning is an emerging technology used to understand the structure of data and fit that data into models that can be used for future predictions. The crop yield prediction performance of machine learning models may be not only dependent on the models, but also dependent on the parameters in data set chosen for training the learning models. Agriculture is the backbone of the Indian economy. Crop yield prediction is an important agricultural problem. In our we propose a model which focuses on predicting the crop yield in advance by analyzing factors like district (assuming same weather and soil parameters in a particular district), state, season, crop type using various supervised machine learning techniques. This helps the farmers to know the crop yield in advance to plan and choose a crop that would give a better yield.

Keywords: Agriculture, Machine Learning, Crop Yield Prediction, Farming, Crops.

### I. INTRODUCTION

Agriculture is the basic of life for many people in India especially in rural Areas. Agriculture along with poultry and dairy farming is practiced as they get food and other raw materials for living and for industries. In India, agriculture can bedated back to Indus valley civilization, with its joint sectors, is one of the largest livelihood in India, especially in rural areas of India. Also India ranks second when it comes to agriculture. In terms of net cropped area, India comes first and then US and China comes after India. India is an agriculturally developed country. Agriculturally improved countries have a better growth in their country's economy. Development in sector of agriculture is very essential for the development of economy of the country. However, many farmers lack in using new advanced technologies and techniques in growing the crops.

Predicting the yield of any desired crop usingmachine learning will make a favour to farmers. Different Machine Learning techniques like Random Forest Regressor and Decision Tree Regressor are utilized for prediction of the crop yield. Input attributes used for prediction are StateName, Season, Crop and Area. Others make use of different climatic factors like amount of sunlight, percentage of rainfall and different agricultural components like soil properties, pH level of soil, soil type, fertilizers required. Disadvantage with the normal procedure is considered climatic factors and soil properties may vary when unexpected natural disasters like floods, drought, heavy rainfall occurs. When we collect a dataset and then predict the crop yield by applying different Machine Learning techniques, the prediction will be most accurate. This project simplifies the entire process of data set collection and making use of appropriate algorithms in prediction of crop yield. With this project, farmers will get an idea to grow which crops in their land to get the maximum yield and profit. With this project, economy of a nation can be improved by increasing the percentage of crops production. By implementing new innovative technological ideas in agriculture sectors, huge impact can be created in production of crops.

### II. PROBLEM STATEMENT

Agriculture is one of the main sources of income in India. There is need to improve the sustainability of agriculture with the rate of increase in suicides of farmer due to crop failure and less yield and losses. Hence, it is a significant contribution towards the economic and agriculture welfare of the countries across the world. The problem statement revolves around prediction of yield of crops considering different climatic conditions of India including various attributes . Goal of this project is to help the farmers to choose the suitable crop to grow in order to get the required yield and the profit. Need for the crop yield prediction is very much essential at this point of time for selecting the right crop.



e-ISSN: 2582-5208

International Research Journal of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal) e:03/March-2023 Impact Factor- 7.868 ww

www.irjmets.com

#### Volume:05/Issue:03/March-2023

III. PROPOSED SYSTEM

In this 21<sup>st</sup> century, it is very common to experiment in every sector by implementing new technological techniques. Making use of new techniques simplifies the process and provides thebetter results. The factors like wind, water supply,soil fertility, rainfall changes unexpectedly, when natural disasters occur. This leads to crop failure, reduction in crop production, scarcity of food products and other materials. A single crop failure can cause huge losses to farmers and countries economic growth. So, there is a desperate need for a new system which can predict the rate of production of crop yield accurately.



#### Fig.No.1: Proposed System

In order to eradicate all such problems, we have proposed this new system, in which high yielding crop will be selected by considering most influencing parameters. This system helps the farmers to meet their crop yield production. The chances for failure of crops will be very less. In this proposed system, Machine Learning techniques like Random Forest Regressor ,Gradient Boosting Regressor and Decision Tree Regressor are made used to predict the rate of production of crop yield considering the input parameters like State Name, District Name, Season, Crop year ,Area, Crop, Production.

### IV. DATA SET

The selected data has a sufficient information regarding the production of crops among various states in India over many years. The dataset used for prediction has been collected from:

https://data.world/thatzprem/agriculture-india.

Dataset has 246091 entries and 7 columns in which four are categorical. With the information provided in the data set, the prediction for crop yield will be calculated by using Machine Learning techniques

The attributes considered in the following dataset are: State Name, Area, Crop, Season.

_	(class 'mandas como framo DataEnamo')			
L>	RangeIndex: 246091 entries, 0 to 246090			
	Data columns (total 7 columns):			
	#	Column	Non-Null Count	Dtype
	0	State_Name	246091 non-null	object
	1	District_Name	246091 non-null	object
	2	Crop_Year	246091 non-null	int64
	3	Season	246091 non-null	object
	4	Crop	246091 non-null	object
	5	Area	246091 non-null	float64
	6	Production	242361 non-null	float64
	dtypes: float64(2),		<pre>int64(1), object(4)</pre>	
	memory usage: 13.1+		MB	

#### Fig.No 2: Data Set Details

@International Research Journal of Modernization in Engineering, Technology and Science [2878]



e-ISSN: 2582-5208

International Research Journal of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal) e:03/March-2023 Impact Factor- 7.868 ww

www.irjmets.com

#### Volume:05/Issue:03/March-2023

#### V. METHODOLOGY

In this project, the system makes use of the Machine Learning techniques to predict rate of crop yield. The programming language used is Python as it is widely accepted for new idea implementations in the field of Machine Learning. In this project, collected data set will be uploaded and prediction for crop yield will be generated by applying Machine Learning techniques like Random Forest Regressor, Gradient Boosting Regressor and Decision Tree Regressor.

The result depends on the information present in the collected data set. Accurate the information about the parameters in the collected datasets, better the results will be.



Fig. No 3: Flow Chart of Crop Yield Prediction VI. RESULT



Comparision of r<sup>2</sup> score values of Decision Tree Regression, Random Forest Regression and Gradient Boosting Regression

## VII. CONCLUSION

This project is proposed to deal with the increasing rate of farmer suicides and to help them to grow financially stronger. This project can be used by the farmers to select the most suitable and high yielding crops to grow in their land and to get the yield of their expectation. Appropriate datasets were collected, studied and trained using machine learning tools. This project contributes to the field of agriculture. One of the most important and



e-ISSN: 2582-5208 mology and Science

# International Research Journal of Modernization in Engineering Technology and Science

(Peer-Reviewed, Open Access, Fully Refereed International Journal) Volume:05/Issue:03/March-2023 Impact Factor- 7.868 wv

www.irjmets.com

novel contributions of the project is providing a list of crops with their productions based on the climatic conditions. This new smart way of predicting the crop yield is effective in calculating the accurate results. If this new way is used for cultivating crops, decision making on selection of crops for farmers.

#### VIII. REFERENCES

- [1] T. U. Rehman, M. S. Mahmud, Y. K. Chang, J. Jin, and J. Shin, "Current and future applications of statistical machine learning algorithms for agricultural machine vision systems," Computers and electronics in agriculture, vol. 156, pp. 585–605, 2019.
- [2] S. J. Reshma and A. S. Pillai, "Impact of machine learning and internet of things in agriculture: State of the art," in International Conference on Soft Computing and Pattern Recognition. Springer, 2016, pp. 602–613.
- [3] M. Rahnemoonfar and C. Sheppard, "Deep count: fruit counting based on deep simulated learning," Sensors, vol. 17, no. 4, p. 905, 2017.
- [4] A. X. Wang, C. Tran, N. Desai, D. Lobell, and S. Ermon, "Deep transfer learning for crop yield prediction with remote sensing data," in Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies. ACM, 2018, p. 50.
- [5] F. Haque, A. Abdelgawad, V. Yanambaka, and K. Yelamarthi, "Crop yield analysis using machine learning algorithms," in World Forum on Internet of Things. IEEE, 2020.
- [6] A. Khattab, A. Abdelgawad, and K. Yelmarthi, "Design and implementation of a cloud-based iot scheme for precision agriculture," in 2016 28th International Conference on Microelectronics (ICM). IEEE, 2016, pp. 201–204.
- [7] A. Chlingaryan, S. Sukkarieh, and B. Whelan, "Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review," Computers and Electronics in Agriculture, vol. 151, pp. 61–69, 2018.
- [8] R. Kumar, M. Singh, P. Kumar, and J. Singh, "Crop selection method to maximize crop yield rate using machine learning technique," in 2015 international conference on smart technologies and management for computing, communication, controls, energy and materials (ICSTM). IEEE, 2015, pp. 138–145.
- [9] X. Yang and M. Sun, "A survey on deep learning in crop planting," in IOP Conference Series: Materials Science and Engineering, vol. 490, no. 6. IOP Publishing, 2019, p. 062053.
- [10] B. van Ginneken, "Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning," Radiological physics and technology, vol. 10, no. 1, pp. 23–32, 2017.