

# Crop Yield Prediction Using Machine Learning Techniques

Ashwini I. Patil<sup>1</sup>, Ramesh A. Medar<sup>1</sup>, Vinod Desai<sup>2</sup>

<sup>1</sup>KLS Gogte Institute of Technology, Belagavi, India

<sup>2</sup>Angadi Institute of Technology and Management, Belagavi, India

## ABSTRACT

Today Indian economy depends upon agriculture. More than 70% of the people in India have taken it as a main occupation, day by day for a particular crop; the formers are not getting proper yield as well as profit due to environmental conditions like soil quality, weather, heavy rainfall, drought, seed damages, fertilizers, pesticides. The farmers not able to produce high production, so taking the historical agricultural data records we can predict the crop yield using machine learning techniques like Linear regression, comparative analysis are done with decision tree, KNN algorithms, using these to achieve the high accuracy and model performance is computed.

**Keywords :** Linear Regression, Crop Yield, Decision Trees, KNN.

## I. INTRODUCTION

Agriculture is the main criteria in Indian economy, as India is totally agricultural based country. In agriculture, the modernization should be applied, because of overall development in the field of agriculture. automatically Indian economy also becomes high.so we need to work on improving in the crop production and achieve high profit .so that government can take the actions on the food risk management, policies and nowadays farmers facing so many challenges so get rid of these problems .have to build a good model for predicting the crop yield using machine learning algorithms are linear regression,Descion trees, K nearest neighbor for comparative analysis, so this model to study and achieve good performance.

Our system achieves a machine learning and use the linear regression algorithm for prediction, it is a linear model that behaves a linear relationship between the dependent and independent variables by fitting a

linear equation to given dataset.it s statistical processes. Linear regression is a supervised machine learning algorithms where outputs are continuous and slop is constant. The key point is assuming that connection between input variables and output variables the result should be linear.in this model input variables are known and target variables we need to predict ,thus input variables supports to determine the target variables and also some uncertainties may remain in results. The predicted outcomes support the input variables.

## II. LITERATURE SURVEY

In paper [1] , the authors persist to research the environmental parameters that affect the crop yield and related parameters. Here a multivariate Regression Analysis is applied for the same. A sample of environmental factors considers a period of 10 years. The System is applied to find the relationship between explanatory variables like AR,AUC, FPI and hence the crop yield as a response variable and R2

value clearly shows that, the yield is especially hooked into AR, AUC and FPI are the opposite two factors that are influencing the crop yield. This research is often enhanced by considering other factors like MSP, CPI, WPI so on. And their relationship with crop yield.

In paper [2], the authors centered on the users and expert reviews across three product categories that are sellers, products and experimental products were conducted. Here the bulk of research cited attempted to finalize the consequences of a user reviews on a product cost and the probability of a purchase. The results of this work help illuminate the contradictory findings across the discrete research study.

In paper [3], the authors have compared feed forward neural networks with traditional statistical methods through linear regression. This work presents the capability of neural networks and their statistical counterparts used in the world of crop yield prediction.

In paper [4], the authors have done the comparison between OLS regression model and special autoregressive model for crop yield prediction in Iowa. The special autoregressive model has shown enormous enhancement in the model performance over the OLS model. The model can provide better prediction than the OLS model and has capability of adjust with the special autocorrelation, which is not considered by the OLS model. This work has shown that NDVI and precipitation are the most important predictors for corn yield in Iowa.

In paper [5], the authors have presented an android based application and an internet site that uses Machine learning methods to predict the foremost profitable crop in the current weather and soil conditions and with current environmental conditions. This system helps the farmer with a sort of option for the crops that will be cultivated, which will be helping them over the long run.

### III. SYSTEM ARCHITECTURE

The crop yield prediction is a method to achieve a high yield of the crop using previous available data like crop name, season, area, production, soil parameters continuously involves all features that used for high yield of the crop. Selection of features which are necessary for target feature. Some are not precisely consider as a yield. Additionally analysis play an important role in the prediction, linear regression approach having two factors response and explanatory variables. Here input parameters acts as independent and the way of predicting variables are dependent variables.

Least square method is capable of fitting each set of points to form polynomial curve and minimizes the residuals. Mostly preferred approach is linear regression model that estimates the value of external domain knowledge and is considered as extrapolation. The input parameters of predicting yield computed by linear regression model within the range of prediction procedure. Farmers also done harvest prediction using their knowledge for a selected field and crop variety. The architecture system combines all small parts and completes the purposed work. This work will proposes a system that processing methods to predicts analyzed agricultural datasets. Our agriculture dataset consists of crop, crop year, seasons, area, crop production.

### IV. METHODOLOGY

The proposed system has the following steps for crop yield prediction using linear regression method. Predicting the analyzed data .and which is prediction entity depends on input entities.

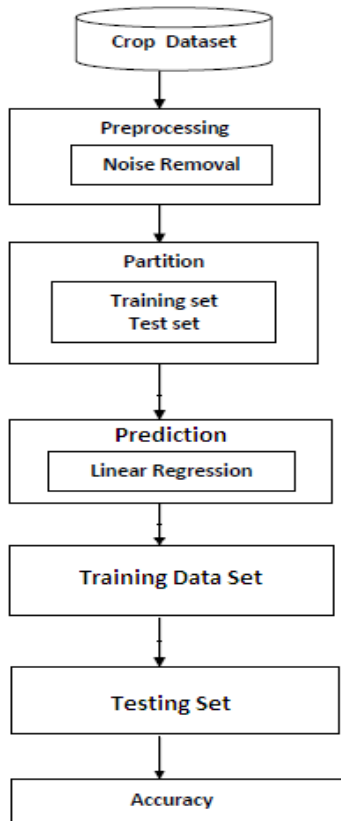


Figure 1 : Linear Regression model

The main steps are as below:

The first step is to input an experimental information set, may be the crop information and soil information along with their outcomes. Few techniques will gather, format and organize the information but only the raw information is scarce to figure with the model. Second step is to collect the data and then pre-process the dataset for the junk or noise removal.

Third step is to separate the data into training and testing datasets: the information must be partitioned into two datasets. The training dataset is having the greatest rate of information, so it will train most of the example datasets to produce the yield. The training samples are collected and the testing set will use the remaining information to see the system performance.

The fourth step is to apply the linear regression on the trained datasets: here the system depends on the complexity of data and also the structure as per the

need. Accordingly the event modeling and the structures are adjusted.

The next step is to apply the trained linear regression model on the test datasets and calculate the R values, and then compare the values with the unique models of linear regression models.

This is one of the simple and effective models for crop yield prediction.

### V. PERFORMANCE AND METRICS

In recent analysis, effectiveness associated with carries with its techniques with categories of effectiveness, Precision, accuracy of the model through the optimization LR proposed model.

Analyzed crop accuracy compared with LR- Proposed, Decision Tree and K-NN

Table 1 : Crop Accuracy Comparison

Algorithm	Cotton	Sugarcane
LR- proposed	99	98
Decision Trees	43	42
K-NN	46	45

The point plot of linear regression is shown in figure.

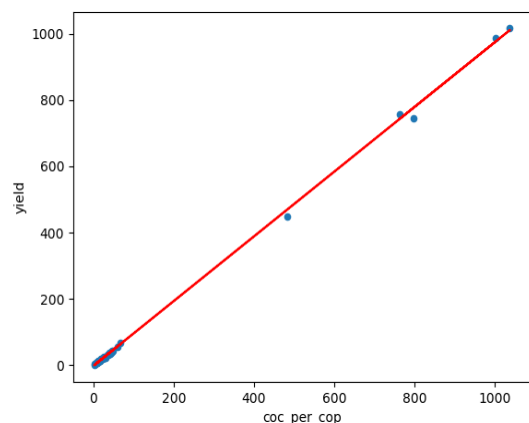


Figure 2 : Point plot of Linear Regression.

## VI. CONCLUSION

Accurate crop yield prediction across different districts helps the farmers in getting the better profit in India. The yield estimation models are used in the agriculture to increase the yield production to satisfy the demands. The readings are used for model inputs and the linear regression algorithms offered convincing estimation accuracy though higher prognostic power that could even be reformulated using alternative crop assessments in longer periods. To improve the system for greater yield and field utilization for crop importance.

## VII. REFERENCES

- [1]. V. Sellam and E. Poovammal, "Prediction of Crop Yield using Regression Analysis," IEEE Trans. Knowl. Data Eng., vol.23, no. 10, pp. 1498-1512. Sept.2010.
- [2]. D Ramesh, B Vishnu Vardhan, "Analysis of Crop Yield Prediction using Data Mining Techniques," International Journal of Research in Engineering and Technology (IJRET), Vol.4, 2015.
- [3]. Raju Prasad Paswan, Shahin Ara Begum, "Regression and Neural Networks Models for Prediction of Crop Production," International Journal of Scientific & Engineering Research, Volume 4, Issue 9, ISSN 2229-5518, September 2013.
- [4]. Li Zhang, Liping Lei, and Dongmei Yan, "Comparison Of Two Regression Models For Predicting Crop Yield," Published by the IEEE IGARSS 2010.
- [5]. Prof. D.S. Zingade, Omkar Buchade, Nilesh Mehta, Shubham Ghodekar, Chandan Mehta, "Machine Learning-based Crop Prediction System Using Multi-Linear Regression," International Journal of Emerging Technology and Computer Science(IJETCS), Vol 3, Issue 2, April 2018.
- [6]. Rajshekhar Borate, "Applying Data Mining Techniques to Predict Annual Yield of Major Crops and Recommend Planting Different Crops in Different Districts in India," International Journal of Novel Research in Computer Science and Software Engineering, Vol. 3, Issue 1, pp: (34-37), April 2016.
- [7]. Dr. Rakesh Poonia, Sonia Bhargava, "Prediction of Crops Methodology using Data Mining Techniques," International Journal of Innovative Research in Science, Engineering and Technology, Vol. 6, Issue 10, October 2017.
- [8]. Suraparaju V, Misra B, Singh CD, "Machine learning approach for forecasting crop yield based on climatic parameters," International Conference on Computer Communication and Informatics (ICCCI-2014); Coimbatore, India. 2014 Jan 3-5.
- [9]. Dakshayini Patil, Dr. M. S. Shirdhonkar, "Rice Crop Yield Prediction using Data Mining Techniques: An Overview," International Journal of Advanced Research in Computer Science and Software Engineering, Volume 7, Issue 5, ISSN: 2277 128Xv, 2017.
- [10]. Ramesh A. Medar and Vijay. S. Rajpurohit, "A Survey of data mining techniques for crop yield prediction," IJARCSMS, Volume 2, Issue 9, September 2014 pg. 59-64.
- [11]. Veenadhari, S., Bharat Misra, D Singh, "Data mining Techniques for Predicting Crop Productivity - A review article," IJCST, International Journal of Computer Science and technology march 2011.

### Cite this article as :

Ashwini I. Patil, Ramesh A. Medar, Vinod Desai, "Crop Yield Prediction Using Machine Learning Techniques", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 7 Issue 3, pp. 312-315, May-June 2020. Available at doi : <https://doi.org/10.32628/IJSRSET20736>  
Journal URL : <http://ijsrset.com/IJSRSET20736>