

ML Methods for Crop Yield Prediction and Estimation: An Exploration



M. Alagurajan, C. Vijayakumaran

Abstract: Machine learning Has performed a essential position within the estimation of crop yield for both farmers and consumers of the products. Machine learning techniques learn from data set related to the environment on which the estimations and estimation are to be made and the outcome of the learning process are used by farmers for corrective measures for yield optimization. This paper we explore various ML techniques utilized in crop yield estimation and provide the detailed analysis of accuracy of the techniques.

Keywords: Machine Learning, Crop Yield,

I. INTRODUCTION

Agriculture is an important ingredient to mankind as it's the major source of livelihood [1]. The world population is significantly increasing that makes the monitoring and estimation of crop production a necessity [9]. To ensure food security and economic growth of a country, agriculture has to be monitored and optimized. The goal of any country is to achieve Most yield rate of crop the usage of fewresources [2]. The estimation yield is tough it depends on different factors like, temperature, moderate rainfall, humidity, soil, crop cultivar, weather, cultivation practices, shape, texture and color of the crop, surface texture and so on. The aim of is to explore some of the ML methods utilized in estimation of crop yield and the parameters used to make the estimations. This paper also aims at implementing some of the machine learning techniques to make estimations on some selected crops. remaining organized as follows: phase 2 presents literatureassessment, segment 3 explore some machine learning techniques used in making estimations, section 4 presents the experiments and result and finally section 5 present the conclusion part of the paper.

II. LITERATURE REVIEW

In [2], authors proposed a crop selection technique that aims to solve the problem of Crop choice improves yield fee selected crop specific season then finally maximize financial boom of the country. In [3], authors use Support Vector Machine technique to classify growing plant as weeds or crop. In their research they frequently capture the image of the plants and perform some feature extractions to identify weeds that prevent the actual crop plant from growing. Authors in this paper use features like color, shape and texture for the classification and statistical textural features.

In [4], authors proposed an architecture for estimation of rice yield called rice yield estimation system (RYPS).

The proposed system in this research predicts rice yield by taking the weather condition of an environment as attributes for the model to make estimations. The first step in this research is data acquisition, then data segmentation to correlate with the objectives and then multiple linear regression, AdaBoost, support vector machine regression and nonlinear regression are applied on the training-set to determine values of parameters.

In [5], authors review machine learning techniques used to make estimations on crop yield. Authors apply sequential minimal optimization using WEKA data mining apparatus on a data-set obtained from 27 localities of Maharashtra. Criteria favor in this research includes: precipitation, less temperature, moderate temperature, high temperature and reference crop evapotranspiration.

In [6], authors used ANN to crop harvest estimation. This research, back propagation method was used to train DNN with three invisible layers to calculate overall cost of the output.

In [7], authors predict two classes called good yield and bad yield based on some environmental features like moderate humidity, moderate temperature and overall rainfall. In this research a hybrid model was used to optimize the features and the features undergoes preprocessing, feature selection and Gray Wolf Optimizer along with Support Vector Machine to obtain optimal accuracy.

In [8], authors used deep neural networks to predict soybean yield estimation. The yield used in this research was equipped from USDA national agricultural statistics service information Stat apparatus between 2003 and 2016 and the satellite data used in the research was collected from NASA's MODIS land.

In [9], authors perform analysis on groundnut data for past 8 years using some machine learning techniques. The techniques used multiple linear regressions, ANN, regression tree, K-Nearest Neighbor. Estimation was done based on environmental features.

Revised Manuscript Received on February 05, 2020.

* Correspondence Author

M. Alagurajan, Department of CSE, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India.

C. Vijayakumaran, Associate Professor, Department CSE, SRM Institute of Science and Technology, Chennai, Tamil Nadu, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Table 1: Overview of Different Estimation Techniques

Title	Method Proposed/Used	Parameters	% Accuracy
Crop choosing technique to increase Crop harvest Rate using ML Technique[2]	Crops Selection Method	Quality of crop, crop hybridization, fertilizer, urea, potash	Depends on the predicted value
SVM Classification Scheme of Maize Crop [3]	Support Vector Machine	Shape, Texture, Color	82%
Rice Yield Estimation Model Using Data Mining [4]	Rice Yield Estimation System (RYPS)	Rainfall, Temperature, Humidity	NILL
Rice Crop harvest Estimation in India using SVM [5]	Support Vector Machine	Less warmth, moderate warmth, high warmth and hint crop dehydration	78.76%
A DNN Approach for Crop choosing and harvest Estimation in Bangladesh [6]	ANN, SVM, Logit Regression and decision Forest	high and less warmth, common drizzle, dankness, climate kinds of area sorts of alchemical fertilizer, sorts of clay, sand structure, clay balance, clay moisture, clay rational, clay response and clay character.	Accuracy Varies for different technique
Big Analytics for Crop estimation using optimization technique [7]	SVM, SVM_GWO	Moderate temperature, moderate humidity, general rainfall and production yield	SVM (72.47%) SVM_GWO(77.09%)
Deep Neural Networks For Crop Yield Estimation [8]	3D Convolution Neural Network	Surface temperature	NILL
Peanut Crop harvest Estimation Using ML Techniques[9]	Multiple linear Regression, Regression Tree, KNN and ANN.	Soil, Environmental and Abiotic attributes	Accuracy Varies for different technique

Nadu. The parameters considered are moderate rainfall, humidity, urea, temperature and soil.

III. CROP YIELD ESTIMATION METHODS

A. Artificial Neural Network

An ANN is a network chain weighted processing units that gets input from the old unit or another unit outside and forwards the output of its processing to the subsequent unit. A neural network may have two layers (input and output) or multiple layers. The number of layers determines the accuracy of the estimation. It is a topological device learning approach that often uses multi-layered perceptron and lower back-propagation set of rules.

B. SVM

The intention of aid vector device is to gain a non linear feature the usage of kernel function. Support vector regression is the SVM used in estimation crop yield. The widely used radial basis functions are kernel and polynomial function.

C. Decision Forest

The decision forest technique is based on tree assembly. In this method, multiple trees of randomly sub sampled features are generated to form a forest and the moderate value of the estimation of specific tree is taking for the evaluation of the output.

IV. EXPERIMENTS AND RESULTS

In this paper we implement ANN, SVM and decision forest technique on our dataset to determine the accuracy of the techniques. The techniques were used to predict the yield production of rice, potatoes and wheat in Chennai of Tamil

Table 2: Evaluation of Rice

Method	Training (%)	Testing (%)	Accuracy (%)
ANN	70	30	96.4
SVM	70	30	73.3
RF	70	30	90.7

Table 3: Evaluation Measures of Potatoes

Method	Training (%)	Testing (%)	Accuracy (%)
ANN	70	30	96.1
SVM	70	30	65.2
RF	70	30	88.7

Table: 4 Evaluation Measures of Wheat

Method	Training (%)	Testing (%)	Accuracy (%)
ANN	70	30	96.4
SVM	70	30	66.3
RF	70	30	90.3

V. CONCLUSION

In this paper, we have explored some techniques used in the estimation of crop yield for promoting food security and economic growth of individual country. Some techniques like decision forest, SVM and ANN were evaluated using some dataset based on some parameters and the accuracy of the methods was presented under the result section of this paper.

REFERENCES

1. R. B. Guruprasad, K. Saurav, and S. Randhawa, "Machine Learning Methodologies for Paddy Yield Estimation in India: a Case Study," *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019.
2. R. Kumar, M. Singh, P. Kumar, and J. Singh, "Crop Selection Method to maximize crop yield rate using machine learning technique," *2015 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM)*, 2015.
3. S. S. Athani and C. Tejeshwar, "Support Vector Machine-Based Classification Scheme of Maize Crop," *2017 IEEE 7th International Advance Computing Conference (IACC)*, 2017.
4. U. K. Dey, A. H. Masud, and M. N. Uddin, "Rice yield estimation model using data mining," *2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 2017.
5. N. Gandhi, L. J. Armstrong, O. Petkar, and A. K. Tripathy, "Rice crop yield estimation in India using support vector machines," *2016 13th International Joint Conference on Computer Science and Software Engineering (IJCSE)*, 2016.
6. T. Islam, T. A. Chisty, and A. Chakrabarty, "A Deep Neural Network Approach for Crop Selection and Yield Estimation in Bangladesh," *2018 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, 2018.
7. S. Sharma, G. Rathee, and H. Saini, "Big Data Analytics for Crop Estimation Mode Using Optimization Technique," *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, 2018.
8. A. S. Terliksiz and D. T. Altlylar, "Use Of Deep Neural Networks For Crop Yield Estimation: A Case Study Of Soybean Yield in Lauderdale County, Alabama, USA," *2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, 2019.
9. V. shah and prachi shah, "Crop Yield and Rainfall Estimation in Tumakuru District using Machine Learning," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, pp. 1094–1097, Jan. 2019.

AUTHORS PROFILE



M. Alagurajan pursuing M Tech department of CSE, SRMIST, Chennai, Tamil Nadu, India.



C. Vijayakumaran, Associate professor, Department of CSE, SRMIST, Chennai, Tamil Nadu, India.