iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.36058

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI June 2021- Available at www.ijraset.com

Review of Crop Yield Prediction using Machine Learning Techniques

Kale Jaydeep Narayan
Computer Engineering Department, Sanjivani College of Engineering Kopargaon

Abstract: Machine learning (ML) could be a helpful decision-making tool for predicting crop yields, in addition as for deciding what crops to plant and what to try throughout the crop's growth season. To help agricultural yield prediction studies, variety of machine learning techniques are used. I performed a literature review (LR) to extract and synthesize the algorithms and options employed in crop production prediction analysis. Temperature, rainfall, and soil types are most common measure used in the prediction as per my knowledge, whereas Artificial Neural Networks is the foremost normally used methodology in these models.

I. INTRODUCTION

ML technologies are employed in a variety of fields like social media, transportation, image processing, speech recognition, self driving cars, stock market trading and many more. Agriculture has been using machine learning for several years. Crop yield prediction is one of precision agriculture's most difficult problems, and numerous models have been suggested and confirmed so far. Because crop production is affected by a range of parameters such as climate, weather, soil type, fertiliser use, and seed variety, this challenge necessitates the use of many datasets. This suggests that predicting agricultural yields is not a simple operation rather, it entails a series of complex stages.

Crop yield prediction methods can now reasonably approximate the actual yield. But better yield prediction action is still desired. ML is a strategy that can deliver superior yield prediction depending on numerous parameters. ML may uncover knowledge from datasets by identifying patterns and connections. The models must be trained with datasets that depict the outcomes depending on previous happening. The predictive model is developed by making use of a variety of features. The model parameters are set during the training phase using previous data.

The available data is divided into two parts. First part is used for training phase and remaining data in second part is used for training in testing phase. When attempting to create a high-performance prediction model, ML studies provide a variety of problems. It's critical to choose the correct algorithms to tackle the problem at hand, and the algorithms and underlying platforms must be able to handle large amounts of data.

I conducted a literature analysis to acquire an overview of what has been done on the application of machine learning in crop yield prediction (LR). A LR identifies possible research gaps in a certain problem area and provides guidance to practitioners and scholars interested in conducting additional research in that area. This research compiles all of the existing literature on the application of machine learning to crop production prediction.

II. LITERATURE SURVEY

P. Charoen-Ung et al. designed a system for predicting sugarcane yield. The author used the data collected from set of sugarcane plots around a sugar mill in Thailand and considered features like plot characteristics, sugarcane characteristics, the plot cultivation schemes, rain volume in their study. They performed prediction using Random forest algorithm. Finally the author concluded that their results were better than non machine learning baselines (i.e. non technical methods) [1].

A. Shah et al. proposed a framework for predicting the crop yield per acre. They used the yield and weather data collected from United States Department of Agriculture. They basically worked on Corn yield data. The humidity, yield, temperature, rainfall were the parameters used by the author in their study. The multivariate polynomial regression, support vector machine regression and random forest models were proposed by the author for the crop yield prediction. Finally they found that SVM is better than other two algorithms[2].

To forecast Maize yield a prediction model was proposed by I. Ahmad et al. They used Crop Management data, Climate data, Soil data, Satellite images of Land cover for predicting yield. They implemented SVM, Decision tree, Random forest, K-nearest neighbour algorithm as a part of their study and developed two models one is Crop model and second one is remote sensing model. The author concluded that remote sensing could be utilised for yield forecasting due to the increased advantages of a smaller input dataset, and if the goal is to measure specific stress and the interaction of plant genetics with soil and environmental variables, a crop model is a highly valuable tool[3].

©IJRASET: All Rights are Reserved

M. Kalimuthu et al. implemented a system for predicting which crop can be cultivated depending on current weather condition in order to have good yield production. Temperature, humidity, moisture content, Seed type were the features used by the author in their study. As a result, for the given input parameter, the seed is projected as an output. According to the author their work may be of great assistance to those farmers who are in need of assistance in anticipating crops in order to establish a sustainable future. They used Naive Bayes algorithm for making the decision of which crop can be cultivated[4].

J. Torres-Tello et al. proposed yield prediction model in aeroponics. Based on environmental variables that may be manipulated and/or measured in the production system, this work proposed a yield prediction model for aeroponic crops in a controlled environment. The Dense Neural Network (DNN), Random Forest based on decision trees (RF) and Support Vector Regression (SVR) were implemented for the study. The author found that DNN was better than other two. The author used features like Temperature, humidity, CO2, pH, Dissolved solids for yield prediction. The author laid out the major steps to take in order to develop a reliable yield prediction (and automated manufacturing) tool that requires less human interaction and yields a higher profit margin [5].

R. Cunha et al. implemented crop yield prediction model using crop calendars, remote sensing data and weather forecast information. The Artificial Neural Network (ANN) was the algorithm used by them in their study. The author used weather and soil data instead of huge satellite photos, which is computationally cheaper to analyse. This differed from past work, which used satellite data for direct farm observation for yield prediction. The system proposed by the author can be used to predict yield before planting, whereas NDVI-based forecasts can only be used after plants have reached a particular growth stage. Their findings suggested that agricultural stakeholders can gain insight into prospective output even before planting, which can aid decision-making in farm and government operations such as machinery leasing, contracting, pricing negotiation, and logistics planning. In the case of noisy yield local data, their system delivered accurate results even with lower data requirements for varied crops. The author concluded that their model is scalable enough to provide accurate predictions for a variety of crops around the world assuming credible local yield data is available.[6].

R. Singh et al. developed a system for detecting most suitable condition for plant growth. The research was performed on Marigold plant. The humidity, temperature, soil temperature and moisture and light intensity on the plant growth was the parameters monitored using IoT based monitoring system. They implemented Logistic Regression, Gradient Boosting Classifier and Linear Support Vector Classifier (SVC) algorithms as a part of their study. The author discovered the most favourable settings for the marigold plant using data visualisation, and there trained model can then be used to determine the rate of growth of the marigold plant just by providing it with the physical environment[7].

R. N. Bhimanpallewar et al. developed agriculture advisory system. The inputs to their system were Soil and environmental parameters while output was crop specific suitability level. They considered 4 types of suitability level like S1- Suitable, S2- Moderately suitable, S3- Marginally suitable and N1- Not suitable. They used hybrid machine learning algorithm which worked better than existing supervised classification techniques. Existing supervised classification techniques are outperformed by a mixed machine learning system[8].

S. S. Kale et al. performed crop production prediction using ANN algorithm. The dataset used in their study was obtained from Indian Government website. The author concluded that due to complexity and severity of crop parameters a linear methodology is insufficient to conclude relationship between factors and crop yield. Traditional linear regression can be replaced with Artificial Neural Network methods to give better accuracy for crop prediction[9].

M. B. Villanueva et al. proposed melon yield prediction model. They used leaves for their study to classify the crop into 6 categories. They implemented ANN s a part of their study[10].

A. Crane-Droesch used Deep Neural networks for yield prediction from weather conditions. They worked on the yield data on corn crop obtained from US Midwest[11].

H. Zhong et al. developed a system for identifying seed varieties for maximizing crop yield. They studied how soil, weather attributes affect crop yield. Machine learning with a weather forecasting procedure used to accurately predict the yield of different seed varieties in a given site[12]..

D. Tedesco-Oliveira et al. predicted cotton yield from color images acquired by simple mobile device using Linear regression algorithm. Tousif Osman et al. performed prediction of Crop Production by analysing Prior Environment Factors like Temperature, Rainfall, Humidity, Sunshine, Cloud coverage, Wind speed, crop area, crop productions. They implemented Linear Regression, Neural network as a part of their study. They found that Neural networks performance was better than performance of linear regression[13].

The table I summarizes the most commonly used crops, features and algorithms in study of crop yield prediction.

©IJRASET: All Rights are Reserved

TABLE I
COMMONLY USED CROPS, FEATURES AND ALGORITHMS

Crops	Features	Algorithms
Sugarcane, Maize, Marigold	Humidity, yield, temperature, rainfall, moisture content, Soil attributes	Random Forest, SVM, Naive Bayes DNN, ANN, Logistic Regression, Linear regression

III.CONCLUSION

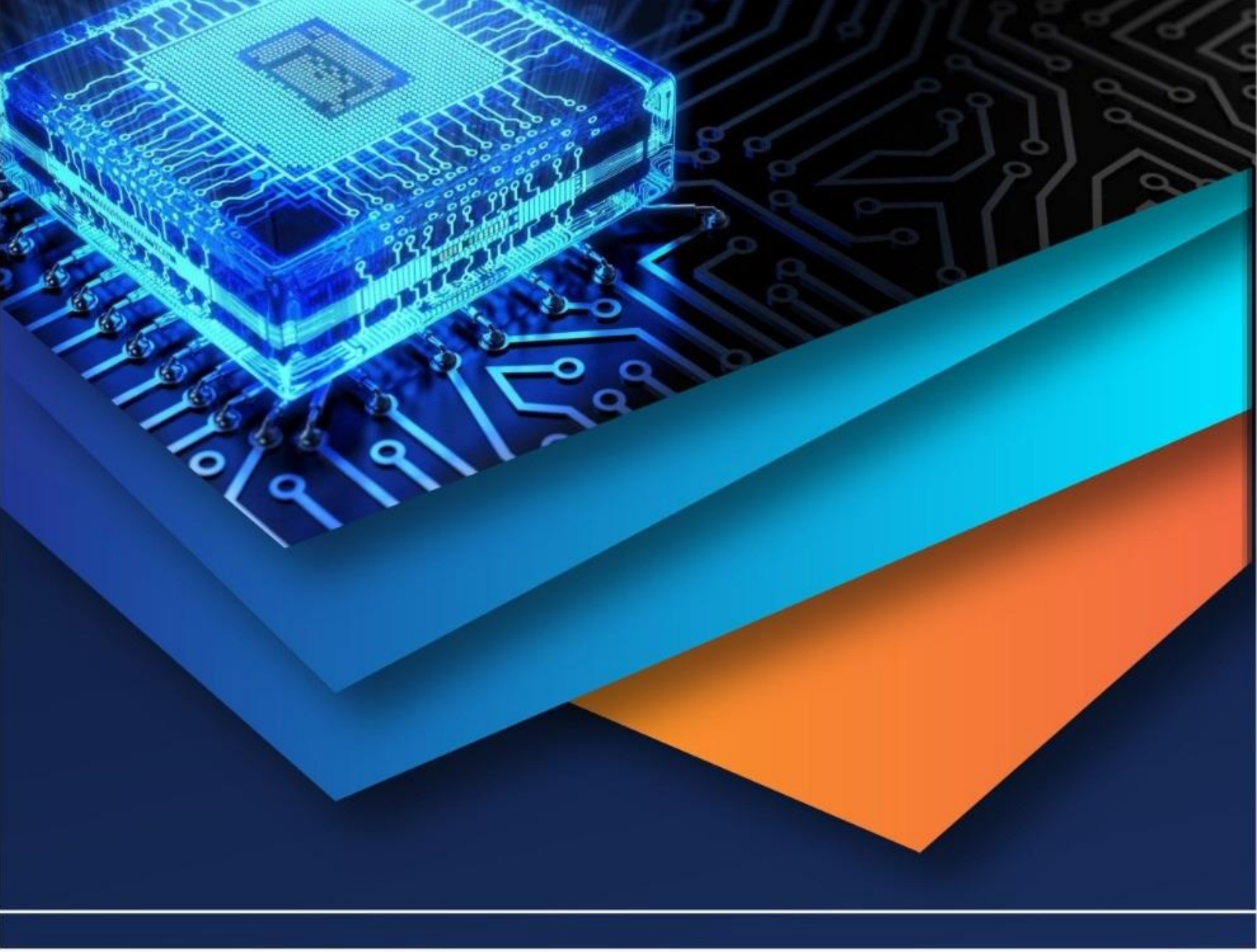
This research revealed that, depending on the scope of the research and the availability of data, the selected publications employ a number of elements. Each paper looks into yield prediction using machine learning, however the features vary. The scale, geological location, and crop of the research all varies. The features chosen are determined by the dataset's availability. According to studies, models with more characteristics may not always deliver the highest yield prediction performance. Several algorithms have been employed in various research.

The findings reveal that no definitive conclusion can be taken about which model is the best, but they do show that some machine learning models are utilised more frequently than others. The random forest, neural networks, linear regression, and gradient boosting tree are the most commonly employed models. I believe that this work will open the road for more research into the agricultural yield prediction problem's development.

REFERENCES

[1] P. Charoen-Ung and P.Mittrapiyanuruk, “Sugarcane yield grade prediction using random forest with forward feature selection and hyper-parameter tuning”, Advances in Intelligent Systems and Computing, vol. 769, pp. 33–42. (2019)
[2] A. Shah, A. Dubey, V. Hemnani, D. Gala and D. R. Kalbande, “Smart Farming System: Crop Yield Prediction Using Regression Techniques”, Proceedings of International Conference on Wireless, 2018
[3] I. Ahmad, U. Saeed, M. Fahad, A. Ullah, M. Habib ur Rahman, A. Ahmad, J. Judg,” Yield Forecasting of Spring Maize Using Remote Sensing and Crop Modeling in Faisalabad-Punjab Pakistan”, Journal of the Indian Society of Remote Sensing (October 2018).
[4] M.Kalimuthu, P.Vaishnavi, M.Kishore “Crop Prediction using Machine Learning”, IEEE, 2020
[5] J. Torres-Tello, S. Venkatachalam, L. Moreno and Seok-Bum Ko ,“Ensemble Learning for Improving Generalization in Aeroponics Yield Prediction”,IEEE, 2020
[6] R. Luiz de Freitas Cunha, B. Silva, “ESTIMATING CROP YIELDS WITH REMOTE SENSING AND DEEP LEARNING”, 2020
[7] R. Singh, S. Srivastava, R. Mishra, “AI and IoT Based Monitoring System for Increasing the Yield in Crop Production”,IEEE, 2020
[8] R. N. Bhimanpallewar and M. R. Narasingarao, “Alternative approaches of Machine Learning for Agriculture Advisory System”, IEEE, 2020
[9] S. S. Kale, P. S. Patil, “A Machine Learning Approach to Predict Crop Yield and Success Rate”, IEEE, 2019
[10] M. B. Villanueva, Ma. Louella M. Salenga, “Bitter Melon Crop Yield Prediction using Machine Learning Algorithm”, (IJACSA), Vol. 9, No. 3, 2018
[11] A. Crane-Droesch, “Machine learning methods for crop yield prediction and climate change impact assessment in agriculture”, Environ. Res. Lett. 13 (2018)
[12] H. Zhong, X. Li1, D. Lobell, S. Ermon, Margaret L. Brandeau, “Hierarchical Modelling of Seed Variety Yields and Decision Making for Future Planting Plans”, 2017
[13] D. Tedesco-Oliveira R. Pereira da Silva, W. Maldonado Jr., C. Zerbato , “Convolutional neural networks in predicting cotton yield from images of commercial fields”, Computer and Ele. In Agri. 171 (2020) https://doi.org/10.1016/j.compag.2020.105307

©IJRASET: All Rights are Reserved



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)