

## **Literature survey for estimate the crop yield using data analytics**

**S. Kulkarni, S. N. Mandal, G. S. Sharma, M. R. Mundada and Meeradevi, "Predictive Analysis to Improve Crop Yield using a Neural Network Model," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018.**

As the industry that feeds the nation's people and contributes to GDP, agriculture has historically been of utmost importance. Crop output fluctuates depending on a number of variables, such as soil characteristics, climate, elevation, and irrigation method. The estimation of the yield based on this mutual dependency of the aforementioned elements has not kept pace with technological advancements. In order to analyse and forecast crop production over seasons in various districts, a data-driven model that learns from historical soil as well as rainfall data has been built. This study is focused on rice as a specific crop. The intended hybrid neural network model determines the best combinations of soil characteristics and combines them with the pattern of rainfall in a chosen area to evolve the anticipated crop production. The Time-Series technique in Supervised Learning serves as the foundation for the predictive analytic model for rainfall. Recurrent Neural Networks, another subset of machine learning, are the technology utilised to make the ultimate prediction of crop yield. The final predictions obtained were successful in illustrating the interdependence between soil factors for yield and weather features thanks to two cooperating data-driven models operating at the backend.

**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.**

In-season crop yield estimation can be used for a variety of purposes, including helping farmers enhance production, optimising the supply-demand cycle for fertilisers, pesticides, and other agricultural products, predicting prices, and calculating the risk levels for agricultural insurance. Due to its ability to be applied to a wide range of crops, growing environments, and environmental circumstances, crop yield estimating models based on remote sensing satellite data and weather data have gained widespread adoption. In this

research, we give a case study of yield estimation modelling for paddy crop based on meteorological and soil data at various spatial resolution (SR) levels, namely, at the taluk (finer SR) and district (coarser SR) levels in India. We offer a thorough examination of the yield estimation models' accuracy using various feature sets and machine learning (ML) methodologies. Additionally, we disaggregate district yield data by utilising machine learning models that were developed using yield predictions at the taluk level. The disaggregation of district level data used to forecast taluk level yield has an average error of 6% and a maximum error of 25%.

**S. Nagini, T. V. R. Kanth and B. V. Kiranmayee, "Agriculture yield prediction using predictive analytic techniques," 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), 2016.**

The expansion of agricultural yields and the related agro sector products are the main drivers of India's economy. The most challenging duty for agricultural departments worldwide is predicting agricultural productivity. Numerous factors affect the output in agriculture. The majority of agriculture growth, especially in nations like India, depends on unexpected rainwater. Growth in the agricultural sector is influenced by a number of factors, including water, nitrogen, weather, soil characteristics, crop rotation, soil moisture, surface temperature, and rainfall, among others. Numerous exploratory data analyses and different predictive models were developed for our paper. Additionally, different regression models, including linear, multiple linear, and non-linear models, are examined for their ability to accurately estimate or forecast the agricultural production for a variety of crops in the states of Andhra Pradesh and Telangana.

**S. Sahu, M. Chawla and N. Khare, "An efficient analysis of crop yield prediction using Hadoop framework based on random forest approach," 2017 International Conference on Computing, Communication and Automation (ICCCA), 2017.**

The development of information technology has made big data a hot topic. Agriculture is crucial to human survival and must make a significant contribution in the area of crop data analysis. This study explains how to use a big data technique to find experiences from accurate agricultural information. In this way, effectively acquiring the valuable data

pushes a framework toward significant computational hurdles in crop analysis where data is collected remotely. We plan to use the Hadoop framework for our work in order to store a massive amount of crop data for the storage purpose of enormous data availability in agriculture. In order to increase production, this work provides a better forecast for farmers regarding the crops they should plant in their fields depending on the soil's composition. Hadoop's MapReduce programming model incorporates the random forest technique.

**F. F. Haque, A. Abdelgawad, V. P. Yanambaka and K. Yelamarthi, "Crop Yield Analysis Using Machine Learning Algorithms," 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), 2020.**

Not only does agriculture play a significant role in the expanding economy, but it is also vital to our survival. It is difficult to predict agricultural output since it depends on a variety of factors, including water, ultraviolet (UV) radiation, pesticides, fertiliser, and the amount of land that is covered in that region. Two distinct Machine Learning (ML) techniques are suggested in this paper to analyse crop yield. Support Vector Regression (SVR) and Linear Regression (LR) are two techniques that are well suited for verifying the variable parameters in the prediction of continuous variables using the 140 data points that were collected. The elements listed above have a significant impact on crop output. Mean Square Error (MSE) and Coefficient of Determination ( $R^2$ ), with MSE being around 0.005 and  $R^2$  being approximately 0.85, were used to compute the error rate.

**N. Johnson, M. B. Santosh Kumar and T. Dhannia, "A survey on Deep Learning Architectures for effective Crop Data Analytics," 2021 International Conference on Advances in Computing and Communications (ICACC), 2021.**

In several fields, deep learning has become a precise tool for both image- and non-image-based data analytics. For crop data analytics, smart farming is a significant area where deep learning techniques have been successfully employed. Crop monitoring is a crucial part of smart farming, which aims to increase agricultural productivity. The key problem is making accurate crop yield predictions when there is little information available about environmental conditions. Frequently, crop diseases are the main cause of production loss. Rural farmers in poor nations lack the tools necessary for real-time crop disease

identification, which prevents them from taking appropriate action. Deep learning techniques have recently shown to be more efficient in producing outcomes in crop data analytics that are more accurate. Research on deep learning-based models used in crop yield estimation and crop disease detection is reviewed in this article.

**P. S. Bharathi, V. Amudha, G. Ramkumar, T. J. Nagalakshmi, N. Nalini and P. Jagadeesh, "An Experimental Analysis of Crop Yield Prediction using Modified Deep Learning Strategy," 2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), 2022,**

Agriculture is the most major and crucial component of the nation's economy, and in comparison to other nations, Indian civilization heavily depends on this sector. To maintain the agricultural field in a good state, many climate parameters such as rainfall, temperature, humidity levels, pesticide problems, and so on must be continuously checked. There are many technologies available today that use artificial intelligence to predict weather patterns and provide accurate information to the appropriate user. A revolutionary deep learning approach is presented in this paper to assist agricultural fields in accurately predicting crop output levels. The proposed learning scheme is referred to as the Modified Deep Learning Strategy (MDLS). This MDLS is derived from the K-Nearest Neighbor and Decision Tree Algorithms, two common learning frameworks. The suggested method takes into account factors like pesticide use, rainfall ratios, and temperature levels as prediction limits when examining agricultural yield characteristics. The resulting section clearly illustrates using graphical representations the correct efficiency ratio for each of the algorithms presented.