International Journal of Innovative Research in Engineering
Volume 3, Issue 6 (November-December 2022), PP: 180-182.
www.theijire.com
ISSN No: 2582-8746

Crop Yield Prediction

Lilavati S. Mhaske[1], Mrunal K. Pathak[2]
[1,2] Department of Information Technology, AISSMS IOIT, SPPU, India.

How to cite this paper:
Lilavati S. Mhaske[1], Mrunal K. Pathak[2]. "Crop Yield Prediction.", IJIRE-V3I06-180-182.

Abstract: *India is an agriculture dependent country and the economic status of the country is totally and partially dependent on it. Agricultural yield is affected by organic, economic and seasonal causes. Estimating agricultural production is a very challenging task for this country with regard to the state of the population. In recent days, people growing these products and such products are very unstable in production due to sudden weather effects on the environment and lack of underground water resources. The main goal is to collect data that can be stored and analyzed to predict crop yield. Machine learning techniques are implemented for crop yield prediction. Machine learning (ML) is an essential perspective to obtain real and operational solutions to crop yield problems. From a given set of predictors, ML can predict the goal/outcome using supervised learning. To get the desired outputs, you need to generate a suitable function using a set of variables that will map the input variable to the target output. Crop yield prediction involves predicting crop yield from past historical data that includes factors such as temperature, humidity, ph, rainfall and crop name. It gives us an idea of the best predicted crop to be grown under field weather conditions. These predictions can be made using a machine learning algorithm called Random Forest. It achieves crop yield prediction with the best accuracy value. A random forest algorithm is used to obtain the best crop yield model considering the least number of models. It is very useful to predict crop yield in agriculture sector.*

Key Word*: Forecasting crop yield, prediction of crop yield, crop yield issues, predictiors, machine learning algorithm, supervised learning, Random Forest, Agriculture Sector.*

I. INTRODUCTION

Agriculture is an important sector for the Indian economy and the future of humanity. It also contributes a large part of employment. As time passes, the need for production increases exponentially. In order to mass produce, people use technology in the wrong way. New types of hybrid varieties are created day after day. However, these varieties do not provide the essential content of a naturally produced crop. These unnatural techniques spoil the soil. Most of these unnatural techniques are used to avoid losses, but when the producers of these crops know the accurate information about the yield of the crop, the loss is minimized. To achieve this research is carried out. The information is provided using past weather information, temperature and several other factors. Data mining is a machine learning tool used to look at data in all possible ways and analyze it. After analyzing the data, it is used to make predictions for future purposes. It can be used in several areas. These patterns provide information about the crop.

1. Aim and Objective of the work

Aim

The aim of the project is to predict crop yield using past historical data which includes factors such as temperature, humidity, Ph, rainfall, crop name. It gives us an idea for the finest predicted crop yield which will be cultivate in the field weather conditions.

Objectives

To implement an algorithm for the prediction of crop yield which will be cultivated in the field weather conditions. It gives us an idea for the finest predicted crop yield which will be cultivate in the field weather conditions.

2. Literature Review

A machine learning algorithm is used to predict crop yield. International Review of Innovations in Technological Scientific Research. This paper aims to predict crop yield using Random Forest algorithm based on existing data. Real data from Tamil Nadu was used to construct the models and the models were checked using samples. The random forest algorithm can be used to accurately predict crop yield. This article provides a detailed overview of research on the implementation of machine learning in agricultural production systems. Along with digitization, methods, approaches and advanced computing, machine learning (ML) has emerged, creating new opportunities to discover, measure and evaluate data-intensive processes in agricultural operational sectors. The Random Forest machine learning algorithm provides the best yield prediction accuracy. A sequential model where a simple recurrent neural network performs better in precipitation forecasting, while LSTM is good for

temperature forecasting. This document combines factors such as rainfall, temperature, season, area, etc. to predict yield. The results show that Random Forest is the best classifier when all parameters are combined. A random forest algorithm builds decision trees on different data samples and then predicts the data from each subset and then votes to give the system a better answer. Random Forest used the bagging method to train the data. To increase accuracy, the injected randomness must minimize correlation while maintaining performance. Characteristics such as rainfall, temperature and season were taken into account when predicting crop yield. Other machine learning algorithms are not applied to the dataset Due to the absence of other algorithms, there was a lack of comparison and quantification, so it was not possible to provide a suitable algorithm. Different machine learning techniques that can be used in different areas of forecasting. However, their work does not implement any algorithms and therefore cannot provide a clear insight into the practicality of the proposed work, concluding that machine learning algorithms can predict the goal/outcome using supervised learning. This Paper focuses on supervised machine learning techniques to crop yield prediction.To get the specified outputs, it needs to generate an appropriate function using a set of variables that can map an input variable to a target output. The paper informs about it predictions can be made using the Random Forest ML algorithm, which achieves the most accurate crop prediction when considering the least number of models.

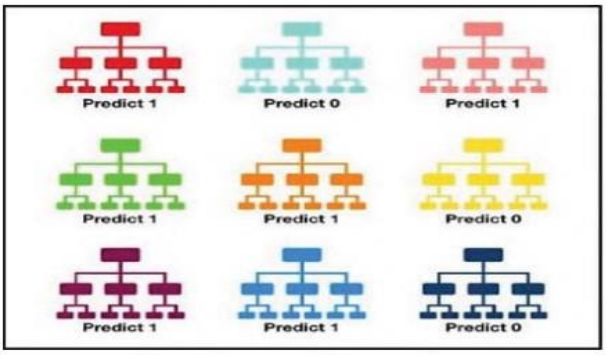
3. Data Acquisition

The following attributes must be given to Dataset-

- Parameters of soil
- Soil Type
- Soil value of ph.
- Climate Parameters
- Wind Rainfall Humidity Temperature
- The cultivation expenses
- Manufacturing

II.PROCEDURE METHODOLOGY

A previously developed country relies on agriculture for its economic development. As the population of a country increases, dependence on agriculture often increases and the subsequent economic process of the country is affected. In this case, the rate of crop yields plays a major role in the economic development of the country. There is therefore a need to increase crop yields. Some biological approaches (e.g. crop seed quality, crop hybridization, strong pesticides) and several chemical approaches (e.g. use of fertilizers, urea, potash) are used to solve this problem. In addition to these approaches, a crop sequencing technique is needed to increase the yield rate of the crop network during the season. One of the current systems we have defined is the Crop Selection Method (CSM) for seasonal realization of net crop yield. We used the example of CSM to show how it enables farmers to produce a higher yield. a) Seasonal crops. During the season, crops can be planted throughout the season. (b) Crop week: crops are also grown during the year. Oh, vegetables, paddy soup, a tour, for starters.

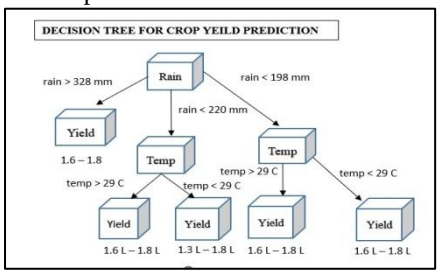


(Fig.1) Tally: Six 1s and Three 0s

It displays crop sequences with the cumulative rate of yield over the season. Method CSM, shown. The process of combining multiple classifiers to solve an in-ensemble learning will boost the web yield rate of crops using the limited land resource and also increase the land's reusability. Basically, the method of crop selection uses a technique where different crop sets are recommended over the years for the same region. There are different choices available for farmers to choose from. In all the decisions, They will pick one and track the outcomes. As output for that region, the mix that is able to offer high yield for the same area is produced. In this way, the CSM approach aims to predict the required crops for the region in question. Farming systems are used strategically in India, depending on the locations where they are most appropriate. Agriculture, organic farming, and commercial farming are the agricultural systems that greatly follow India's agriculture.

Decision Trees- In a decision tree, for predicting the class of the given dataset, the algorithm starts from the root node of the tree.This algorithm compares the values of root attribute with the record (real dataset) attribute and, based on the comparison, follows the branch and jumps to the next node.For the next node, the algorithm again compares the attribute value with the

other sub-nodes and move further. It continues the process until it reaches the leaf node of the tree.



(Fig.2) Decision Tree for Crop Yield prediction

III.COMPARISON RESULT

Dataset consists of few important attributes such as temperature, rainfall, humidity, ph., applied random forest algorithm for classification and regression tasks. Tried to train the model with decision trees but found that random forest algorithm reduced the overfitting problem and also improved accuracy and used Random forest and random forest got more accuracy. The dataset which used is imported from Kaggle repository.From the dataset, used 80% of data for training the model and 20% of data for testing the results to obtain better results and trained the model by applying random forest algorithm.Then, compared the predicted result with the original data set. Later, estimated the accuracy of the model using test samples.Likewise, predicted the accuracy of the model with different algorithms. Then finally concluded that random forest algorithm gives us more accuracy. Hence, used RANDOM FOREST ALGORITHM to train the model.

Random Forest -

At training situation multitude decision trees are made and the output will be divided based on number of classes, i.e. classification, prediction of class. The number of trees is proportional to accuracy in prediction. The dataset includes factors like rainfall, perception, temperature and production. These factors in dataset is used for training. Only two third of the dataset is considered, Remaining dataset is used for experimental basis. The algorithm random forest has 3 parameters – Node tree which describes the n number of trees which need to grow. m try - mentions how many variables need to be taken at a node split. Node size - In terminal nodes it suggest us the number of observation need to take.

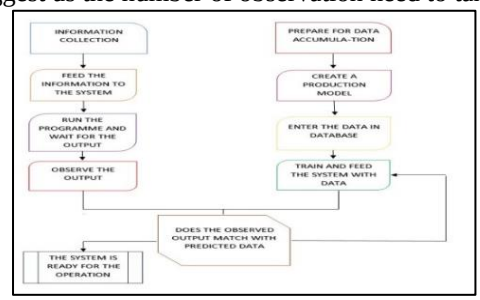


Fig (3) Flowchart for Random Forest Model for Crop Prediction

paired *t*-test was used to determine the difference between baseline and 2 years after regarding biochemistry parameters, and this was confirmed by the Wilcox on test which was a nonparametric test that compares two paired groups. Chi-square and Fisher exact tests were performed to test for differences in proportions of categorical variables between two or more groups. The level $P < 0.05$ was considered as the cutoff value or significance.

IV.CONCLUSION

This paper focuses on crop prediction and yield calculation using machine learning techniques. Several machine learning methods are used to calculate accuracy. The Random Forest classifier was used to predict the yield for the selected district. Implemented a crop prediction system from past data collection. The proposed technique helps farmers in deciding which crop to grow in the field. This work serves to find the acquired knowledge about the crop that can be used for efficient and useful harvesting. Accurate forecast of various specified crops in various districts will help farmers in Kerala. This improves our Indian economy by maximizing the yield of crop production.

References

1. *Namgiri Suresh, N.V.K.Ramesh, Syed Inthiyaz, "Crop Yield Prediction Using Random Forest Algorithm", 2021 7th International Conference on Advanced Computing & Communication Systems (ICACCS) .*
2. *Kiran Moraye , Aruna Pavate , Suyog Nikam and Smit Thakkar4, "Crop Yield Prediction Using Random Forest Algorithm for Major Cities in Maharashtra State", International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ISSN: 2347-5552, Volume-9, Issue-2, March 2021.*
3. *Potnuru Sai Nishant , Pinapa Sai Venkat , Bollu Lakshmi Avinash , B. Jabbe "Crop Crop Yield Prediction based on Indian Agriculture using Machine Learning", 2020 International Conference for Emerging Technology (INCET) Belgaum, India. Jun 5-7, 2020.*
4. *Reema Methew, Jinsu Mani "Crop Yield Prediction Using Machine Learning Algorithm", IJERT 2021.*