Journal of Analysis and Computation (JAC)
(An International Peer Reviewed Journal), www.ijaconline.com, ISSN 0973-2861
Volume XVIII, Issue II, July-December 2024

CROP YIELD PREDICTION USING ML TECHNIQUES

[1]Naveen Kumar Kedia, [2]Saurabh Kumar, [3]Ankit Kumar, [4]Saroj

[1]Assistant Professor, Department of Information Technology, JECRC College
[2]B.Tech Student, Department of Information Technology, JECRC College
[3]B.Tech Student, Department of Information Technology, JECRC College
[4]B.Tech Student, Department of Information Technology, JECRC College

ABSTRACT

Looking at the current situations that Indian farmers are dealing with, we have noticed that there have been a lot of suicides in India over a long period of time. The cause of these suicides is the frequent changes in the Indian Government system and the weather. Farmers may lack knowledge of the crops that are most suited to their soil's quality, nutrients, and composition. This project that seeks to assisting farmers in evaluating the soil's quality to ensure a decent crop production. Before cultivating in an agricultural field, the farmer might use the prediction to estimate the crop's production. A key strategy for finding a workable and efficient solution to this issue is machine learning. The results aren't particularly accurate given the current method, which includes manual counting, climate-smart pest management, and satellite photography. The major goal of this research is to use several machine learning approaches to forecast the agricultural production. Among the classifier models utilized here, Random Forest offers the highest accuracy, followed by Logistic Regression and Nave Bayes. By taking into account variables like temperature, rainfall, area,and other characteristics, the predictions provided by machine learning algorithms will assist farmers in choosing which crop to cultivate to induce the greatest yield. This ties the technology and agricultural sectors together.

Keywords - Agriculture, Machine Learning, Prediction, and Crop Yield.

[1] INTRODUCTION

Since its inception, agriculture has been the main activity in every society and civilization that has existed throughout human history. It is not only a huge part of the expanding economy, but it is also crucial to our survival. It is also a vital sector for the Indian economy and the future of humanity. Additionally, it makes up a sizable amount of

employment. As time goes on, the demand for production has dramatically expanded. People use technology in an utterly incorrect manner in order to produce in large quantities. Every day, new hybrid varieties are created. These kinds, however, don't offer the same critical components as naturally grown crops. These artificial methods degrade the soil.

All of this causes more environmental deterioration. Consequently, this initiative suggests a strategy to forecast the crop's output. Before cultivating the field, the farmer will measure the crop production per acre. Farmers are ignorant about the crops that are most suited to their soil's quality, nutrients, and composition. The initiative suggests using a data mining approach to assist farmers in assessing the quality of the soil. According to the soil type, the algorithm predicts the crop that may be grown there and maximizes crop yield by suggesting the right fertilizer.

[2] LITERATURE REVIEW

In [1] employing a machine learning algorithm to predict agricultural yield. Engineering Science Research Technology International Journal.

This research employs the Random Forest algorithm to forecast the agricultural yield from the available data. The models were constructed using actual data from Tamil Nadu, and they were tested using samples. Crop yield predictions can be made with accuracy using the Random Forest algorithm.

In [2] Random forests for global and regional crop yield prediction. PLoS ONE Journal. Because of its great accuracy and precision, simplicity of usage, and utility in data analysis, RF is a useful and adaptable machine-learning method for agricultural production projections at regional and global scales. The most effective method, Random Forest, outperforms Multiple Linear Regression (MLR).

In [3]. Crop production Ensemble Machine Learning model for prediction. International Journal of Computer Science and Software Engineering (IJCSSE).

AdaNaive and AdaSVM are the suggested ensemble models in this study that will be used to forecast crop production over a given time frame. AdaSVM and AdaNaive were used in the implementation.

SVM and Naive Bayes algorithm efficiency is improved with AdaBoost.

In [4]. Machine learning approach for forecasting crop yield based on parameters of climate. This essay was presented at the International Conference on Computer

Communication and Informatics (ICCCI). In the current study, a user-friendly web page called Crop Advisor was created as a software tool to anticipate the impact of meteorological conditions on crop yields. In a few Madhya Pradesh areas, the C4.5 method is utilised to

determine the climatic characteristic that has the greatest impact on crop yields of particular crops. The decision tree is used to implement the paper.

In[5]. Prediction On Crop Cultivation. International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE) Volume 5, Issue 10, October 2016.

The interpretation of soil test findings and soil analysis are currently done on paper. This has in some way contributed to incorrect interpretation of soil test results, which has led to inaccurate recommendations of crops, soil amendments, and fertilisers tofarmers, resulting in subpar crop yields, micronutrient deficiencies in the soil, and excessive or inadequate fertiliser application. Formulas to Recommend Fertilizer and Match Crops with Soil.

[3] RELATED WORK

Existing Systems: The majority of current systems rely on hardware, which makes them expensive and difficult to maintain. They also don't provide reliable results. Some systems recommend crop sequencing based on market value and yield rate.

System Proposed: The system is designed to address these issues and forecast crops using structured data analysis. Since it is solely a software solution, little consideration of maintenance is permitted. When compared to hardware- based systems, the accuracy would be high since factors like soil composition, soil type, pH value, and others take on a picture throughout the prediction process.

Problem Statement: Using machine learning based on previous crop forecast and soil quality assessments, a high crop yield can be achieved. The primary goal of this project is to forecast crop yield, which will be very helpful to farmers as they plan for the harvest and sale of grain. Implement a machine learning algorithm that provides a more accurate forecast of the best crop for the appropriate region and growing season in our nation. The goal of this study is to forecast yields using location and meteorological information. This study's objective is to examine the prediction of crops that will provide high yield in the specified site taking the climatic and soil conditions into account.

[4] METHODOLOGY

A crucial component of every machine learning system is data. We choose to concentrate on the Indian state of Maharashtra for implementing the system. Data collected at the district level were important since local climates vary. To put the system into place, historical information on the crops and climate of a certain area was required. This information was taken from many official websites. The information on the crops grown in each Maharashtra district was obtained from www.data.gov.in, and the information on the climate was obtained

from www.imd.gov.in. Precipitation, temperature, cloud cover, vapour pressure, and the frequency of rainy days are the climatic factors that have the greatest impact on crop production. Therefore, information on these meteorological variables was acquired on a monthly basis.

Collection of Datasets: During this stage, we gather data from multiple sources and create datasets. And analytics are being used with the provided dataset (descriptive and diagnostic). There are many online sources for abstracts, including Data.gov.in and indiastat.org. The annual abstracts of a crop will be used for at least ten years. These datasets typically permit time series with anarchic behaviour. The primary and necessary abstracts were combined. Global and Regional Crop Yield Predictions Using Random Forests.

Data Partitioning: The entire dataset is divided into two sections, with, for instance, 25% of the data being set aside for model testing and the other 75% being utilised to train the model.

To predict future events Machine Learning Algorithms:

Supervised learning: Using tagged examples, supervised machine learning algorithms may transfer prior knowledge to new data. The system may provide targets for any new input after sufficient training.

The learning algorithm may also distinguish between its outputs and the correct, intended output and detect mistakes in order to adjust the model appropriately. Contrarily, unsupervised machine learning techniques are employed when the data used to train is neither labelled nor categorised. Unsupervised learning examines how systems can extrapolate a function from unlabeled data to describe a hidden structure. The system doesn't determine the correct output in order to characterise hidden structures from unlabeled data, but it may analyse the data and make inferences from datasets.

The most well-known and effective supervised machine learning algorithm, known as random forest, can perform both classification and regression tasks. It works by building a large number of decision trees during training and producing outputs of the class that are the mean prediction (for regression) or mode of the classes (for classification) of the individual trees.

The prediction is more reliable the more trees there are in a forest.

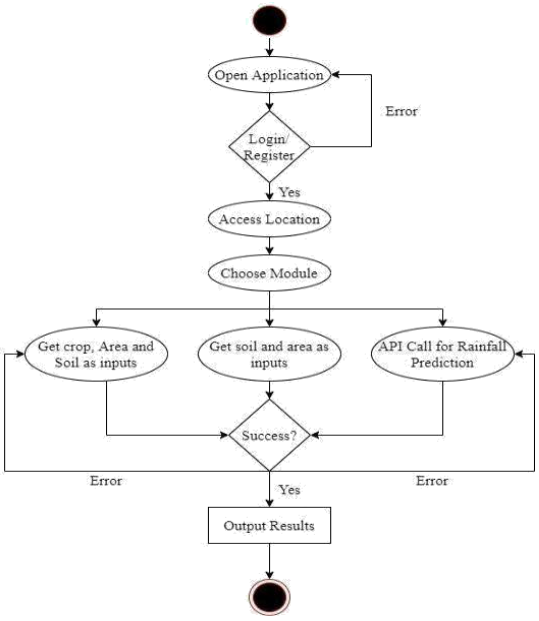

Fig. 4.1: Proposed Approach

[5] CONCLUSION AND FUTURE WORK

The current study demonstrated the possible use of data mining approaches in agricultural production prediction based on the meteorological input factors. The website that was created is easy to use, and all of the crops and study regions that were chosen had predictions that were more than 75% accurate. Any user can utilise the user- friendly website designed for crop yield prediction by submitting climatic data for their preferred crop.

This system contributes to the field of agriculture. One of the most important and novel contributions of the system is suggesting the user the right time to use the fertilizer, this is done by predicting the weather of the next 14 days. Also, the system provides a list of crops with their productions based on the climatic conditions.

The future work is focused on providing the sequence of crops to be grown depending on the soil and weather conditions and to update the datasets time to time to produce accurate predictions. The Future Work targets a fully automated system that will do the same. Another functionality that we are trying to implement is to provide the correct fertilizer for the given crop and location. To implement this through study of fertilizers and their relationship with soil and climate is required. We are also aiming to predict the crisis situation in advance like the recent hike of onion prices.

REFERENCES

[1] P.Priya,U.MuthaiahM.Balamurugan. Predicting yield of the crop using machine learning algorithm. International Journal of Engineering Science ResearchTechnology.
[2] J.Jeong, J.Resop, N.Mueller and team.Random forests for global and regional crop yield prediction.PLoS ONE Journal.
[3]. Narayanan Balkrishnan and Dr. Govindarajan Muthukumarasamy.Crop production Ensemble Machine Learning model for prediction. International Journal of Computer Science and Software Engineering (IJCSSE).
[4]. S.Veenadhari, Dr. Bharat Misra, Dr. CD Singh.Machine learning approach for forecasting crop yield based on climatic parameters. International Conference on Computer Communication and Informatics (ICCCI).