JES
Journal of Engineering Sciences

Vol 13, Issue 02, FEB /2022

ISSN NO:0377-9254

A SURVEY ON CROP YIELD PREDICTION AND EFFICIENT USE OF FERTILIZER

E-mail: mavuram.sailu@gmail.com

Name : Mavuram Sailu

MTECH-CSE

JNTUH College of engineering, Jagtial

Guide:

Email ID: svraju.jntu@gmail.com

Name : Dr.S.Viswanadha Raju

Professor of CSE Department,

JNTUH College of Engineering, Jagtial

ABSTRACT

There are more than a half of India's people who depend on agriculture for their daily necessities. Weather, climate, and other factors in the environment are putting agriculture at risk. Crop yield prediction also relies heavily on machine learning (ML), which helps farmers pick what crops to grow and what they should do during the growing season. CYP features are thoroughly evaluated in this work, and a variety of artificial intelligence methods are used to assess crop production prediction utilising artificial intelligence. The Neural Network has two main drawbacks: reduced relative error and poorer Crop Yield forecast effectiveness. Supervised learning techniques failed to account for the nonlinear link between input and output variables when sorting and grading fruits. Numerous research projects were suggested for agricultural development with the purpose of developing a model that could accurately and efficiently classify crops based on factors such as weather and crop disease, and so on. Analysis of ML methods used to estimate agricultural yields and their accuracy was reported in this study.

1.INTRODUCTION

1.1 MOTIVATION

India has been involved in agriculture in some capacity since the Indus Valley Civilization, making it the country with the most experience in this area. Agriculture and its linked sectors, such as forestry and fisheries, account for 15.4 percent of the GDP (gross domestic product). According to net cropped area, India is the most populous country, with the United States and China in close second and third place. Agricultural production is the country's most important industry, contributing significantly to the overall socio-economic structure of India. Economic expansion in India has resulted in a decline in the agriculture sector's share of the country's gross domestic product. This can be attributed to the industrial revolution.

1.2 PROBLEM DEFINITION

An important question in India's agriculture sector is how to use modern technologies to

get the desired outcomes Rainfall and temperature patterns are getting more unpredictable as new technologies and nonrenewable energy resources are being utilised. Farmers have been unable to precisely predict temperature and rainfall patterns due to the effects of global warming, which has resulted in poorer crop yields. The use of machine learning techniques such as RNN, LSTM, and others can help identify a pattern, which can then be utilised to make accurate predictions and deal with temperature and rainfall variability. The expansion of agriculture in the country will benefit farmers' livelihoods. Agricultural output has been improved in the past by researchers using machine learning.

1.3 OBJECTIVE OF PROJECT

The goal of this research is to employ a variety of machine learning approaches to forecast crop productivity. Mean absolute error (MAE) is used to compare the results of these procedures. Based on characteristics such as temperature, rainfall, acreage, etc., machine learning algorithms can help farmers pick which crops to cultivate in order to maximise production.

The following are the goals that will be pursued in the future:

1. The modulation factor values of ML algorithms vary depending on the different crop feature divisions.

2. ANN is used to reduce the number of input elements. In order to obtain an accurate estimation of crop yield, the best feature was empirically chosen.

When solving complex issues, ML technique regression reduces them to simple linear function optimizations, removing the need for a huge output sample space.

Crop yield estimation can be accomplished using a massive soil dataset and a machine learning system.

A significant increase in crop output was made possible by the ML techniques, which were used to observe agricultural fields.

2.LITERATURE SURVEY

Trio of Priya, Uthaiah, and Balamurugan

Proposed the country's economy is heavily dependent on agriculture.

Farmers face a huge danger from climate and other environmental changes. Practical and effective solutions to this challenge can only be achieved through the use of machine learning (ML). Based on previous data, such as weather parameters, soil parameters, and crop yields, crop yield prediction can be done. This research uses the Random Forest method to anticipate crop yields based on current data. Authentic Tami lnadu data was utilised to develop the models, which were then put through their paces with actual samples. Prior to planting the crop, the farmer will be able to estimate its yield using this prediction tool. Predicting future crop yields with Random Forest is an extremely powerful and popular supervised machine learning approach.

Mishra.s, Mishra.D, and Santra.G.H. H is proposed agricultural crop production is the focus of this work, which aims to reevaluate previous research on the applicability of machine learning approaches. Statistical/Methodological Approaches: This is a brand-new approach to crop management in agriculture. Economic and statistical officials rely on accurate crop production estimates to make key policy decisions like import-export pricing and marketing distribution. One must recognise, however, that these past estimates are not objective estimates, as these estimates

involve a lot of descriptive judgement based on a variety of qualitative characteristics. As a result, realistic crop output predictions based on statistical evidence are urgently needed. Large amounts of data have been generated by this advancement in computers and information storage. Findings: As a result of this challenge, new techniques like machine learning have been developed that may be used to combine the knowledge of the data with the evaluation of agricultural yields. The purpose of this study was to examine these new methodologies so that substantial correlations could be identified between their application to the various variables in the database. Application/Improvement: Artificial neural networks, Information Fuzzy Networks, Decision Trees, Regression Analysis, and Bayesian belief networks are only a few examples of advanced methods available to researchers. Agriculture is a field where time series analysis, Markov chain models, k-means clustering, k nearest neighbours, and support vector machines are used.

Manjula E proposed a new research area in crop production analysis is data mining. In agriculture, predicting crop yields is a major concern. Every farmer wants to know how much he may expect to get out of his crops. In the past, farmers' knowledge of a specific field and crop was used to anticipate yields. In the absence of better data, it is impossible to accurately anticipate yields. When it comes to this, data mining is the way to go. In order to predict crop yields for the following year, several Data Mining techniques are put to the test. Based on past data, this study suggests and implements an algorithm to estimate agricultural yield. By using association rule mining on agricultural data, this can be accomplished. The goal of this study is to develop a crop yield forecast model that can be used in the future. For the district of Tamil Nadu in India, this paper gives a brief investigation of crop yield prediction utilising data mining techniques and association rules. The experimental findings suggest that the proposed approach is able to accurately predict crop yields.

Dahikar, S. S, Rode and S. V proposed by **taking into account numerous climatological processes that affect local weather conditions in different places of the world. Crop yields are directly impacted by these meteorological conditions. The relationship between large-scale climatological occurrences and crop productivity has been studied in a variety of ways. Modeling and prediction using artificial neural networks have been shown to be highly effective. Sensing soil and atmospheric parameters as well as other soil parameters is utilised in crop prediction methodology to choose the best crop to plant. The kind of soil, PH, nitrogen, phosphate, potassium, organic carbon, calcium, magnesium, sulphur, manganese, copper, iron, depth, temperature, rainfall, humidity are some of the variables that might be taken into consideration. We're employing an artificial neural network for that aim (ANN).**

Gonzalo Snchez, Frausto Sols and Ojeda Bustamante are the authors proposed this work for agricultural planning objectives, reliable yield estimation is critical. Practical and effective solutions to this challenge can only be achieved through the use of machine learning (ML). The most accurate yield prediction approach has been the subject of numerous comparisons of ML techniques. Crops and procedures are reviewed far too rarely, and as a result there isn't enough data to help farmers plan their operations. Crop yield prediction is examined using both linear regression approaches and machine learning (ML) in this study. The approaches of multiple linear regression, M5-Prime

regression trees, perceptron multilayer neural networks, support vector regression, and k-nearest neighbour were all compared. The models were validated using four accuracy metrics: RMS, RRSE, MAE, and the correlation factor (a measure of the root mean square error) (R). Real data from an irrigation zone in Mexico was used to create the models. Samples from two consecutive years were used to evaluate the models. Both M5-Prime and KNN have the lowest average MAE error of 18.12% and 19.42%, while M5-Prime also has the lowest RMSE error of 5.14 and 4.91 percent, and KNN has the highest average correlation factor of 0.996%. (0.41 and 0.42). M5-Prime is an excellent tool for large-scale agricultural planning since it produces the lowest number of crop yield models with the lowest error.

SYSTEM ANALYSIS

EXISTING SYSTEM:

- Exiting models mostly used neural networks, random forests, and KNN regression approaches for CYP, as well as various ML techniques for the best prediction.
- There are a number of issues that need to be addressed in the current agricultural yield prediction study utilising machine learning:
- Because of their complexity, ML algorithms are extremely expensive to develop, repair, and maintain.
- To better anticipate crop yield (mustard, wheat), machine learning algorithms merged input and output data.
- With nonlinear data and extreme value data, the regression model was unable to offer a precise prediction because of the linear connection between parameters.
- For yield prediction, current K-NN models were applied, but their performance was dropped due to the nonlinear and highly adaptive difficulties prevalent in KNN models. They worked with a locality model that increased the input vector's dimensionality, resulting in classification problems.
- Because there was a lack of data to estimate crop production, an incorrect decision was made during classification.

Disadvantages:

- As a result of a lack of linearity and adaptability in KNN classification models, yield prediction performance was reduced.

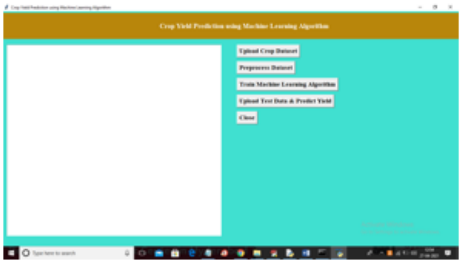
PROPOSED SYSTEM:

- Applied machine learning methods and their quantification are the topic of this research. Rainfall and temperature data are unreliable, thus they are taken into consideration in the work provided here. Rather of focusing on just one factor when trying to predict crop yield, this approach takes into account all of the relevant data.
- CNN, LSTM and DNN were the most commonly utilised algorithms in the trials, although there remained room for development in CYP.

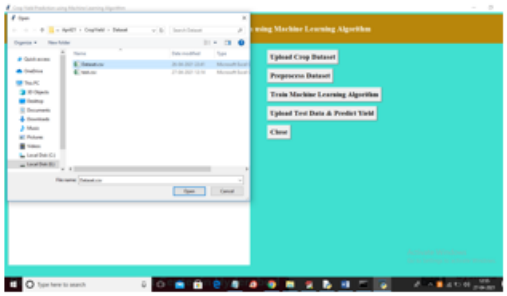
- There are a number of models out there that take temperature and meteorological conditions into account when predicting crop yields.
- To sum up: • The experimental study demonstrated how using machine learning in agriculture can help improve crop forecasting.
- **MODULES:**
- **Upload Crop Dataset**
- • Crop production data is loaded into classification and regression algorithms to predict the crop's name and yield.
- Random Forest Regressor was found to be the most accurate predictor of crop yield in experiments based on Indian government datasets. Predicting rainfall is easier with a sequential model based on a Simple Recurrent Neural Network, but predicting temperature is easier with an LSTM. The yield projection for a specific district can be created by integrating rainfall, temperature, and other criteria like season and area.
-
- Computer-Aided Intelligence
-
- District-by-district estimates of crop yields are the focus of this study. District-by-district and best-yielding crop yields are being forecasted.
-
- Predict Yield by Uploading Test Data
-
- When all parameters are considered, Random Forest is the best classifier. When it comes to determining what crops to cultivate next year, this will benefit farmers and bridge the gap between technology and farming.

SCREENSHOTS

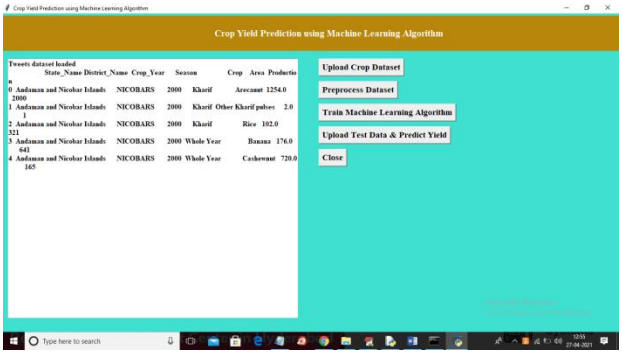
To run project double click on 'run.bat' file to get below screen



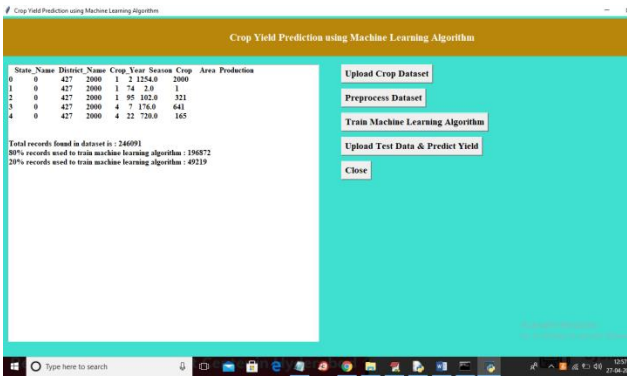
In above screen click on 'Upload Crop Dataset' button to upload dataset



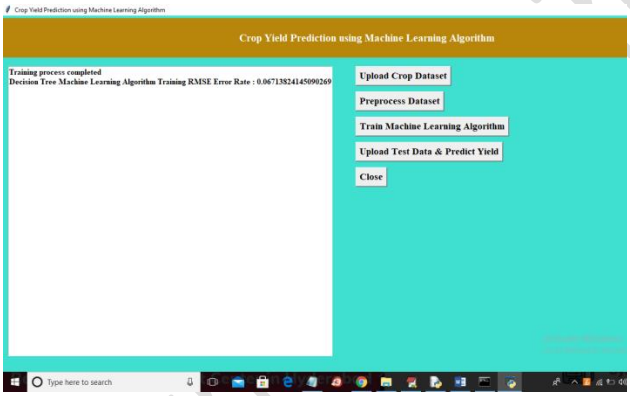
In above screen selecting and uploading 'Dataset.csv' file and then click on 'Open' button to load dataset and to get below screen



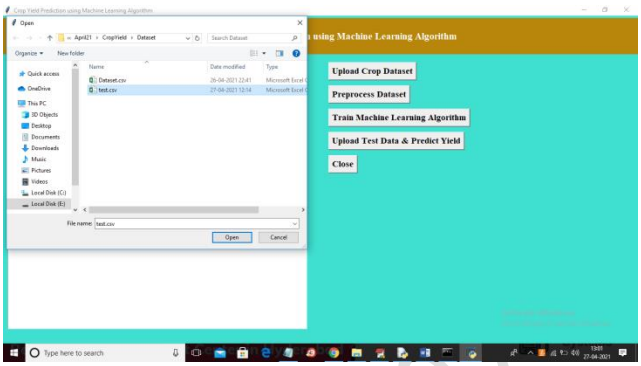
As shown in the screenshot, the dataset contains some non-numeric values. Because ML will not accept non-numeric values, we must preprocess the dataset to assign an ID to each non-numeric value before running ML on it. To begin processing the dataset, click on the 'Preprocess Dataset' button.

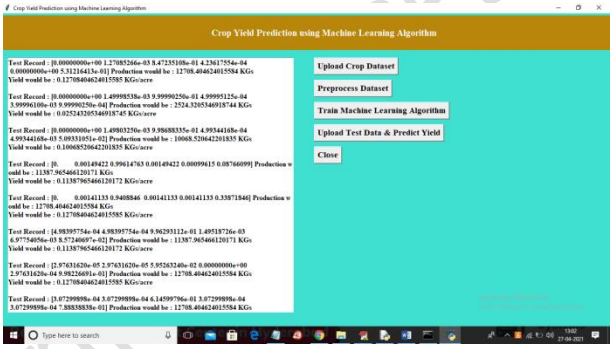


Here, we can observe that the dataset comprises a total of 246091 records, and that the application uses (80%) 196872 records to train ML and uses (20%; root mean square error) 49219 records to test the RMSE (root mean square error) of the ML prediction error rate. Once you've clicked on the 'Train Machine Learning Algorithm' button, you can use the dataset above to train the Decision Tree Machine learning algorithm and compute the prediction error rate.



You can see in this page that machine learning (ML) has been trained and we got a prediction error rate of 0.067% and now the Decision Tree model is complete, so you can input test data and then use the programme to estimate production yield.





Test data values are shown in square brackets, followed by anticipated YIELD per acre, on the above screen. Newlines are used to separate each test record on the screen. As a result, newlines are used to separate each test record and forecast. Using the scroll bar above the text box, you may access all of the data.

CONCLUSION

For the purposes of this study, a wide range of features will be examined, although each of the studies will use ML techniques that are distinct from the features. The availability of data sets influenced the feature selection procedure, which took into account geological position, scale, and crop characteristics. However, adding more features does not necessarily result in better results. Because of this, the number of best-performing traits that were also assessed was reduced. Neural networks, random forests, KNN regression and a range of other ML methods were commonly used in the current models for CYP prediction. The most often used algorithms in the experiments were

CNN, LSTM, and DNN, although CYP still had space for improvement. Pre-existing models have taken temperature and weather conditions into account when making predictions about crop productivity, according to the current study. After all was said and done, the experiment's results showed that using ML to agriculture helped advance crop prediction. However, there was still a lot of opportunity for improvement in the selection of agricultural features influenced by temperature variance. The delay to border topographical areas demanded additional-explicit treatment in following investigations, which were based on these possibilities. To achieve the ideal statistical CO2 fertilisation, deterministic crop models are used instead of machine learning algorithms. If the aforementioned objectives are met, further research on agricultural yield estimation will be improved. Farmers should also take fertiliser into consideration when calculating crop yields in order to make better decisions in the event of underestimated crop yields." In view of the study's findings, we will need to create and develop a CYP model based on DL in the future.

REFERENCES

[1] R. Ghadge, J. Kulkarni, P. More, S. Nene, and R. L. Priya, "Prediction of crop yield using machine learning," Int. Res. J. Eng. Technolgy, vol. 5, 2018.

[2] F. H. Tseng, H. H.Cho, and H. T. Wu, "Applying big data for intelligent agriculture-based crop selection analysis," IEEE Access, vol. 7, pp. 116965-116974, 2019.

[3] A. Suresh, N. Manjunathan, P. Rajesh, and E. Thangadurai, "Crop Yield Prediction Using Linear Support Vector Machine," European Journal of Molecular & Clinical Medicine, vol. 7, no. 6, pp. 2189- 2195, 2020.

[4] M. Alagurajan, and C. Vijayakumaran, "ML Methods for Crop Yield Prediction and Estimation: An Exploration," International Journal of Engineering and Advanced Technology, vol. 9 no. 3, 2020

[5] P. Kumari, S. Rathore, A. Kalamkar, and T. Kambale, "Predicition of Crop Yeild Using SVM Approch with the Facility of E-MART System" Easychair 2020.

[6] S. D. Kumar, S. Esakkirajan, S. Bama, and B. Keerthiveena, "A microcontroller based machine vision approach for tomato grading and sorting using SVM classifier," Microprocessors and Microsystems, vol. 76, pp.103090, 2020

[7] P. Tiwari, and P. Shukla, "Crop yield prediction by modified convolutional neural network and geographical indexes," International Journal of Computer Sciences and Engineering, vol. 6, no. 8, pp. 503-513, 2018.

[8] P. Sivanandhini, and J. Prakash, "Crop Yield Prediction Analysis using Feed Forward and Recurrent Neural Network," International Journal of Innovative Science and Research Technology, vol. 5, no. 5, pp. 1092-1096, 2020.

[9] N. Nandhini, and J. G. Shankar, "Prediction of crop growth using machine learning based on seed," Ictact journal on soft computing, vol. 11, no. 01, 2020

[10] A. A. Alif, I. F. Shukanya, and T. N. Afee, "Crop prediction based on geographical and climatic data using machine learning and deep learning", Doctoral dissertation, BRAC University) 2018.

[11] A. Fuentes, S. Yoon, S. C. Kim, and D. S. Park, "A robust deeplearning-based detector for real-time tomato plant diseases

and pests' recognition," Sensors, vol. 17, no. 9, pp. 2022, 2017.

[12] J. Sun, L, Di, Z. Sun, Y. Shen, and Z. Lai, "County-level soybean yield prediction using deep CNN-LSTM model," Sensors, vol. 19, no. 20, pp. 4363, 2019.

[13] K. A. Shastry, and H. A. Sanjay, "Hybrid prediction strategy to predict agricultural information," Applied Soft Computing, vol. 98, pp. 106811, 2021.

[14] D. A. Bondre, and S. Mahagaonkar, "Prediction of Crop Yield and Fertilizer Recommendation Using Machine Learning Algorithms," International Journal of Engineering Applied Sciences and Technology, vol. 4, no. 5, pp. 371-376, 2019.

[15] B. Devika, and B. Ananthi, "Analysis of crop yield prediction using data mining technique to predict annual yield of major crops," International Research Journal of Engineering and Technology, vol. 5, no.12, pp. 1460-1465, 2018.