AGRICULTURE YIELD PREDICTION USING MACHINE LEARNING

Dr.I.Selvamani[1]., L.Sravani[2]., B.Shruthi Reddy[3]., K.Indu[4]., S.Navya Sree[5]

1 Professor, Department Of ECE., Malla Reddy College Of Engineering For Women., Maisammaguda., Medchal., Ts, India (✉i.selvamani@gmail.com)

2,3,4,5 B.Tech ECE, (19RG1A0492, 19RG1A0465, 19RG1A0483, 19RG1A04A7),

Malla Reddy College Of Engineering For Women., Maisammaguda., Medchal., Ts, India

Abstract— Over the last two decades, most agricultural products in India have been severely impacted by the effects of climate change. Policymakers and farmers would benefit from knowing the crop output in advance so that they can plan for marketing and storage. As a result of this work, farmers will have a better idea of what to expect from their crops before they even start planting them. An interactive prediction system prototype is developed in an effort to address the problem. The machine learning algorithm and an intuitive web-based graphical user interface will be implemented to create this kind of system. The farmer will have access to the forecasting findings. We can anticipate crop production with the use of several algorithms designed specifically for such data analytics in crop prediction. The algorithm is a random forest. While we can analyze the causes of our predicament, we lack the technology and solutions necessary to change the weather, temperature, humidity, rainfall, and moisture. India's agricultural sector offers several opportunities for boosting the country's GDP. Predictions of agricultural yields may also be made with the use of data mining. Data mining, in its broadest sense, is the practice of looking at data from many angles and synthesizing the results. In order to perform both classification and regression, random forests, the most popular and powerful supervised machine learning algorithm, build many decision trees during training and then produce an output class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees**.**

Keywords— Agriculture, Machine Learning, crop-prediction, Supervised Algorithms, Crop yield, Data Mining.

I. INTRODUCTION

Indian agriculture is crucial to the country's economy. The weather has a huge role on India's agricultural output. Rainfall is crucial to the success of rice farms. Helpful suggestions for foreseeing

forecasting future crop productivity and doing an analysis will aid farmers in achieving maximum crop yields. Predicting crop yields is a critical issue in agriculture. Farmers used to be able to anticipate their yield based on the results of the previous year. Consequently, we may anticipate crop production with the aid of these algorithms, which were developed specifically for this sort of data analytics in crop prediction. The algorithm is a random forest. All of these algorithms, and the connections between them, are playing an increasingly important role in agriculture, and their use is expanding. The agricultural sector has been deteriorating since the introduction of cutting-edge tools and methods. As a result of these many innovations, many individuals now devote their time and energy to the production of artificial items, many of which are hybrids that contribute to an undesirable way of living. People now lack traditional agricultural knowledge, such as when and where crops should be planted. Seasonal climatic conditions are also being altered by these cultivating practices, which works against key assets like soil, water, and air and ultimately leads to food insecurity. There is no correct answer and technology to overcome the scenario we are in by analyzing all these concerns and challenges including weather, temperature, and numerous other aspects. India's agricultural sector offers several opportunities for boosting the country's GDP. The quantity and quality of harvests may be increased and enhanced in a variety of ways. Predictions of agricultural yields may also be made with the use of data mining.

The main objectives are

a. To use machine learning techniques to predict cropyield.
b. To provide easy to use User Interface.
c. To increase the accuracy of crop yield prediction.
d. To analyse different climatic parameters (cloud cover,rainfall, temperature)

II. LITERATURE REVIEW

In [1] Using a machine learning system to forecast agricultural output. Journal of Engineering Science and Technological Research International. In this study, we apply the Random Forest method to estimate future crop yields using historical records. The models were developed using actual data from Tamil Nadu, and the results were validated.

examples included. Predicting agricultural yields with the use of the Random Forest Algorithm is possible.

Random forests for predicting agricultural yields at the national and regional levels [2]. Journal of PLoS ONE. Based on the results we've collected, it's clear that RF is a powerful and flexible machine-learning approach for predicting crop yields on both local and global sizes. When compared to multiple linear regression (MLR), the most effective method is Random Forest.

In [3]. Predicting crop yields using an ensemble of machine learning techniques. An International Forum for Software Engineering and Computer Science. In this research, we present an ensemble model to forecast crop yields over time using two individual models: AdaNaive and AdaSVM. AdaSVM and AdaNaive were used for the actual implementation. SVM and the Naive Bayes method both benefit from AdaBoost's enhanced performance.

The use of machine learning to predict agricultural production from climatic variables is shown in [4]. The ICCCI paper presents research presented at the annual conference. In this study, we created a user-friendly website called Crop Advisor that uses weather data to provide

predictions about crop yields.In order to determine which climatic characteristic has the greatest impact on agricultural yields for which crops in which Madhya Pradesh districts, the C4.5 algorithm is utilized. Decision Tree is used to implement the article.

In[5]. Agriculture Forecasting. In October 2016's Volume 5 Issue 10 of the International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE). Soil analysis and the decoding of soil tests are currently carried out manually. Poor crop yields, soil micronutrient deficiencies, and excessive or insufficient fertilizer application can all be traced back to inaccurate interpretation of soil test results, which in turn has led to poor recommendations of crops, soil amendments, and fertilizers to farmers. Crop-Soil Compatibility and Fertilizer Recommendation Formulas.

Crop yield prediction analysis using data mining techniques is presented in [6]. International Journal of Research in Engineering and Technology; hereafter referred to as "IJRET" for short. The primary objective of this work is to provide a user-friendly interface for farmers that provides an analysis of rice production using existing data. Several Data mining methods were utilized to forecast harvest success in order to maximize agricultural production. Such as the K-Means method for predicting atmospheric pollution.

Machine learning applications in agriculture are discussed in [7]. Document Identifier: DOI:10.17485/ijst/2016/v9i38/95032 Indian Journal of Science and Technology, Volume 9 Issue 38, October 2016.

An enhanced unclear cluster analysis is offered using GPS-based color photos for identifying plant types, soil types, and residue types. Many of the factors that affect crop output and the ratio of yield that may be raised during cultivation are discussed in this study.

In [8] In this work, we provide a thorough analysis of studies conducted specifically on the topic of using machine learning in agricultural production systems. New possibilities for understanding, quantifying, and analyzing data-intensive processes in agricultural operating sectors have developed with the advent of machine learning (ML) and other big data technologies, techniques, approaches, and high-performance computers. The paper is put into practice by using Support Vector Machines (SVP).

In [9]. Precision Agriculture from a Drone: Determining Yield for Crop Insurance. Address: 5th & 6th Floor, Artur Centre, Gokhale Cross Road, Model Colony, Pune - 411016 Symbiosis Institute of Geoinformatics, Symbiosis International University. The term "precision agriculture" (PA) refers to the practice of employing geospatial techniques and remote sensors to track changes in a field and adapt farming methods accordingly. Crop stress, irrigation methods, pest and disease prevalence, and other factors may all contribute to growth variation in agricultural fields. Ensemble Learning (EL) is used to implement the paper's methods.

In [10]. Institute on the Environment, University of Minnesota, St. Paul, Minnesota 55108, United States of America. Random Forests for Global and Regional Crop Yield Predictions. The results demonstrate that RF is an innovative and successful machine learning approach for predicting agricultural yields at both regional and global scales. K-Nearest Neighbor, Support Vector Regression (SVG) was used to create this paper's implementation.

III. METHODOLOGY

All Machine Learning Systems rely heavily on data. We settled on India's Maharashtra State as our primary target for the system's rollout. Due to regional variations in weather, district-level statistics were required. The method required archival information on the crop and climate of a certain area. This information was collected from several governmental portals. We accessed www.data.gov.in for information on the agricultural output of each Maharashtra district, and www.imd.gov.in for climatic information. The frequency of rainy days, the amount of precipitation, the average temperature, the amount of cloud cover, and the vapour pressure all have significant impacts on the crop. As a result, monthly measurements were taken of these climate factors.

Gathering of Datasets: At this stage, we are compiling databases and collecting data from numerous sources. Both descriptive and diagnostic analytics may be applied to the input data collection. Data.gov.in and indiastat.org are only two examples of the many websites that provide abstracts. The annual crop abstracts will be utilized for at least 10 years. The chaotic behavior of time series is often tolerated in these data sets. Integrated the most important and obligatory summaries. Predicting Crop Yields at the National and International Level Using Random Forests.

Data Partitioning: The whole dataset is split in half, with one half utilized for model training and the other used for model testing.

forecasting future occurrences Learning Algorithms in Machines: In order to apply what has been learnt in the past to new data, supervised machine learning algorithms need labeled samples of data. Once the system has been adequately trained, it will be able to generate targets for any novel input. The learning algorithm may also compare its findings to the proper, intended output and identify mistakes, allowing it to make necessary adjustments to the model. In contrast, unsupervised machine learning techniques are used in situations when the training data is not itself subject to observation.

without a name or a category. In unlabeled data, unsupervised learning analyzes how a system might infer a function to describe a latent structure. Instead of determining the correct output, the system evaluates the data and may make inferences from datasets in order to identify hidden structures in unlabeled data.

Classification via Random Forests: In order to perform both classification and regression, random forests, the most popular and powerful supervised machine learning algorithm, build a large number of decision trees during training and then produce outputs of the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees. The more trees there are, the more reliable the forecast will be.

IMPLEMENTATION

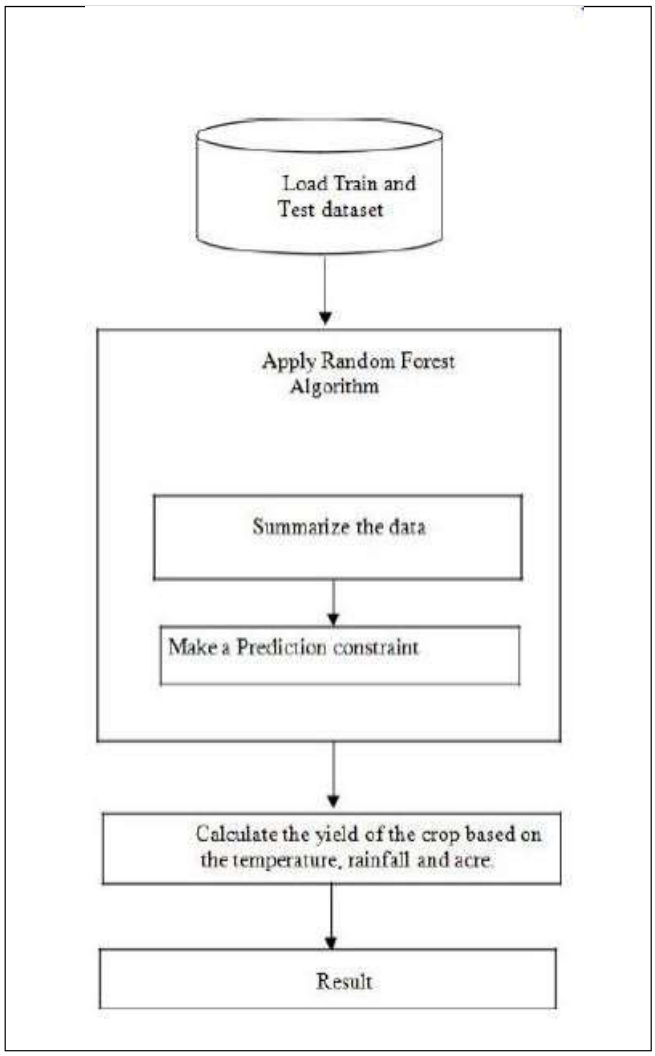

Fig. 1. Proposed Approach

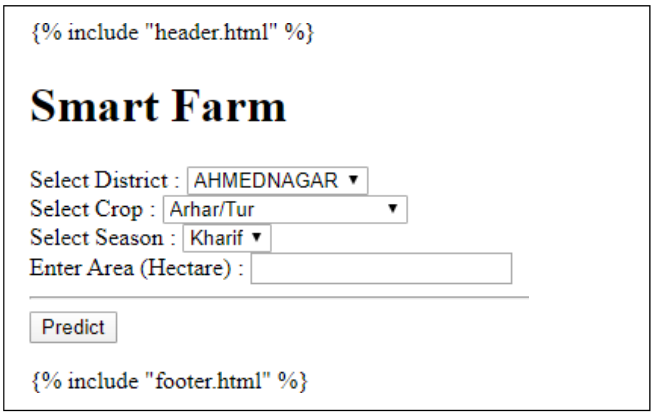

Fig. 2. Home page

Fig. 2 shows the home page of the website where the person accessing the website enters the details such as the district, crop, season and the area in Hectare and by clicking on predict the result is printed

District_Name	Season	Crop	PJan	PFeb	PMar	PApr	PMay	PJ
AHMEDNAGAR	Kharif	Arhar/Tur	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Kharif	Bajra	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Kharif	Gram	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Kharif	Jowar	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Kharif	Maize	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Kharif	Moong(Green (	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Kharif	Pulses total	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Kharif	Ragi	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Kharif	Rice	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Kharif	Sugarcane	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Kharif	Total foodgrain	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Kharif	Urad	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Rabi	Jowar	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Rabi	Maize	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Rabi	Other Rabi pulses	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Rabi	Wheat	3.099	0	1.671	23.129	4.646	
AHMEDNAGAR	Summer	Maize	3.099	0	1.671	23.129	4.646	

Fig. 3. Data set

Fig. 3. It is the snapshot of the final processed data set that is being used for this project

CONCLUSION AND FUTURE SCOPE

This research demonstrated the possible utility of data mining methods in forecasting agricultural production based on meteorological input characteristics. All of the crops and study districts picked had forecast accuracies of 75% or above, and the produced website is easy to use. The user-friendly website created for estimating crop yield may be utilized by any user, regardless of their chosen crop, by just giving the meteorological data of that location.

REFERENCES

[1] U.Muthaiah, P.Priya M.Balamurugan.Using a machine learning system to forecast agricultural output. Journal of Engineering Science and Technological Research International.

[2]. Researchers J.Jeong, J.Resop, and N.Mueller.Predicting agricultural yields at the national and regional levels using random forests.Journal of PLoS ONE.

[3].The names of the authors are Narayanan Balkrishnan and Dr. Govindarajan Muthukumarasamy.Predicting crop yields using an ensemble of machine learning techniques. An International Forum for Software Engineering and Computer Science.

Researchers S.Veenadhari, Bharat Misra, and CD Singh.The use of machine learning to predict agricultural production using weather data. Computer Communications and Informatics (ICCCI) International Conference.

(Reference #5) Shweta K. Shahane and Prajakta V. Tawale.Agriculture Forecasting. October 2016 Volume 5 Issue 10 of the International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE).

D. Ramesh & B. V. Vishnu Vardhan, [6]D. Predicting crop yields: a data mining analysis. The International Journal of Engineering and Technology, or IJRET for short.

[7]Gour Hari Santra, Subhadra Mishra, and Debahuti Mishra. Crop production using machine learning applications in agriculture. DOI:10.17485/ijst/2016/v9i38/95032, Indian Journal of Science and Technology, October 2016. Dionysis Bochtis, Simon Pearson ID, and Konstantinos G. Liakos. Patrizia Busato. Dimitrios Moshou. Agriculture using machine learning. Spearson@lincoln.ac.uk, University of Lincoln, Lincoln Institute of Agri-food Technology (LIAT), Brayford Way, Brayford Pool, Lincoln LN6 7TS, UK.

It was [9] Baisali Ghosh. Precision Agriculture from a Drone: Determining Yield for Crop Insurance. Five and Sixth Floor, Atur Centre, Gokhale Cross Road, Model Colony, Pune 411016, India. Symbiosis Institute of Geoinformatics, Symbiosis International University.

Jig Han Jeong, Jonathan P. Resop, Nathaniel D. Mueller, David H. Fleisher, Kyungdahm Yun, Ethan E. Butler, and Soo-Hyung Kim [10]. Predicting Crop Yields at the National and International Level Using Random Forests. The Environment Institute, University of Minnesota, St. Paul, Minnesota 55108, USA.

[11] (2009). Ecochem Online. Crop yields and the state of the soil. Last updated on Monday, January 28th, 2009.From http://ecochem.com/healthy_soil.html (retrieved on March 4, 2009).

[12] Organization for Food and Agriculture. (2006). Market Conditions for Agricultural Products. 37–39.

"Prediction of crop yield using Regression Technique," Aditya Shastry, H.A. Sanjay, and E.Bhanushree, International Journal of Computing, volume 12, issue 2, pages 96–102, 2017, ISSN:1816-914. A Model for Predicting Crop Yield, E.14]E. Manjula and S. Djodiltachoumy.

Computational Intelligence and Informatics: An International Journal, Volume 6 Issue 4 March 2017

[14] D.Santhiya, S.Nishanthini, and Mrs.K.R.Sri PreethaaK.Vani Shree, "CropYield Prediction," International Journal of Engineering, Technology, and Science, IJETS, Volume III, Issue III, March 2016, ISSN (P): 2349-3968, ISSN (O):2349-3976

[15] Article by Jharna Majumdar, Sneha Naraseeyappa, and Shilpa Ankalaki titled "Application of big data in agriculture through data mining techniques"J Big Data 2017;4:20 DOI: 10.1186/s40537-017-0077-4 Majumdar et al.

[16]In 2015, D. Ramesh and B. Vardhan published "Analysis of crop yield prediction using data mining techniques" in the International Journal of Research in Engineering and Technology, volume 4, issue 1.

[17] To apply data mining methods to the study of agriculture and related disciplines, see Yethiraj N. G.'s article from Volume 1, Number 2 (December 2012).

[18] Zia Zelua, 2009. Agriculture land grading was performed using a Decision Tree in an expert system based on spatial data mining. Intelligent Computing, Technology, and Automation: The Second International Conference.Oct10-11,China