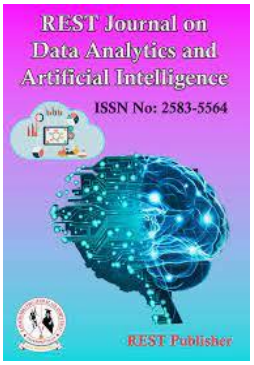
REST Journal on Data Analytics and Artificial Intelligence

Vol: 2(1), March 2023

REST Publisher; ISSN: 2583-5564

Website: http://restpublisher.com/journals/jdaai/

DOI: https://doi.org/10.46632/jdaai/2/1/3



Crop Yield Prediction Using Machine Learning

***P. Kalpana, I. Anusha Prem, S. Josephine Reena Mary, Rev. Sr .ArockiaValan Rani**

2St.Joseph's College of Arts and Science for Women, Hosur, Tamil Nadu, India.

*Corresponding author Email: **kalpana912001@gmail.com**

Abstract: *The majority of India's agricultural products have been negatively impacted by climate change in terms of performance over the past 20 years. Prior to harvest, crop output predictions would aid farmers and policymakers in deciding on the best course of action for marketing and storage. Before cultivating on the agricultural field, this project will assist the farmers in learning the yield of their crop, enabling them to make the best choices. By creating a working prototype of an interactive prediction system, it tries to find a solution. It will be put into practise to implement such a system with a user-friendly web-based graphic user interface and the machine learning algorithm. The farmer will have access to the prediction's outcomes. So, there are various ways or algorithms for this type of data analytics in crop prediction, and we can anticipate crop production with the aid of those algorithms. It employs the random forest algorithm. There are no suitable technologies or solutions to deal with the scenario we are in, despite the analysis of all these concerns and problems, including weather, temperature, humidity, rainfall, and moisture. In India, there are numerous ways to boost agricultural economic development. Data mining can be used to forecast crop yield growth. Data mining is, in general, the process of reviewing data from various angles and distilling it into pertinent information. The most well-known and effective supervised machine learning algorithm, random forest, can perform both classification and regression tasks. It works by building a large number of decision trees during training time and producing output of the class that is the mean prediction (for regression) or mode of the classes (for classification) of the individual trees.*

Key words: *Agriculture, Machine Learning, crop-prediction, Supervised Algorithms, Crop yield, Data Mining.*

1. INTRODUCTION

The Indian economy is based mostly on agriculture. In India, weather has a major impact on agricultural output. Rainfall is mostly necessary for rice cultivation. Timely guidance to anticipate the It is necessary to make an analysis of future agricultural productivity in order to assist farmers in maximizing crop yield. A significant issue in agriculture is yield prediction. Farmers used to forecast their yield based on yield data from previous years. So, there are various ways or algorithms for this type of data analytics in crop prediction, and we can anticipate crop production with the aid of those algorithms. It employs the random forest algorithm. There are an increasing number of applications for big data analytics approaches in agriculture using all these algorithms and the relationships among them. Since the development of new cutting-edge technology and methods, agriculture is slowly deteriorating. Owing to these numerous inventions, individuals are concentrating on creating artificial and hybrid products, which can result in an unhealthy lifestyle. The growing of crops at the proper time and location is not something that modern people are aware of. Due to these cultivating methods, the seasonal climate is also changing at the expense of basic resources like soil, water, and air, which causes food insecurity. There are no suitable solutions or technologies to deal with the situation we are in right now, despite the analysis of all these problems and issues related to weather, temperature, and other factors. In India, there are numerous strategies to boost the economic growth in the sphere of agriculture. There are numerous techniques to boost and improve the crop Output and the quality of the crops. Data mining can be used to forecast crop yield growth. The primary goals are

1. To forecast crop yield using machine learning techniques.
2. To offer a user interface that is simple to use.
3. To improve crop yield predictions' precision.
4. To analyse various climate factors (cloud cover, precipitation, and temperature).

2. LITERATURE REVIEW

In [1] crop yield prediction using a machine learning algorithm. International Journal of Engineering Science, Technology, and Research. This study employs the Random Forest algorithm to forecast the crop's production based on the data already

available. The models were developed using actual data from Tamil Nadu, and samples were utilised to test the results. For precise crop yield forecast, use the Random Forest algorithm. For the prediction of both global and regional crop yields, see [2].Journal PLoS ONE. 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]. Machine learning ensemble model for predicting crop production. Journal of Computer Science and Software Engineering International (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. A machine learning method for predicting agricultural yield based on climate factors is presented in [4]. 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 Adviser 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]. Crop cultivation prediction Volume 5, Issue 10, October 2016, of the International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE). 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 to farmers, resulting in subpar crop yields, micronutrient deficiencies in the soil, and excessive or inadequate fertiliser application. Formulas to Suggest Fertilizer and Match Crops with Soil. Agricultural yield prediction analysis using data mining techniques is presented in [6]. IJRET: The article that was published in the international journal for engineering and technology research. The primary goal of this study is to develop a user-friendly interface for farmers that provides an analysis of rice production based on the data that are already accessible. Several data mining approaches were employed to forecast crop yields in order to maximise crop output. Such as the K- Means algorithm to predict the atmospheric pollution factor Uses of Machine Learning Methods in the Production of Agricultural Crops, [7]. In October 2016, the Indian Journal of Science and Technology published Volume 9(38), An increased indistinct cluster analysis for categorising regions of interest in plants, soil, and residue is offered from GPS- based colour pictures. The study provides a number of parameters that can help improve crop productivity and boost the yield ratio during cultivation. In [8] In this study, we provide a thorough analysis of the literature on the use of machine learning in agricultural production systems. Along with big data technologies, methods, and high- performance computing, machine learning (ML) has arisen to create new potential for unravelling, quantifying, and analysing data-intensive processes in agricultural operational sectors. Support vector machines (SVP) are used to implement the paper. In [9]. An investigation of crop insurance yield utilising precision farming from an aerial platform. 5th and 6th Floor, Artur Center, Gokhale Cross Road, Model Colony, Pune, 411016: Symbiosis Institute of Geoinformatics Symbiosis International University. Precision agriculture (PA) involves the use of remote sensors and geospatial approaches to detect field differences and address them in various ways. In an agricultural field, crop stress, irrigation techniques, the presence of pests and diseases, among other factors, may be the cause of crop growth variability. Ensemble Learning is used to implement the paper (EL). In [10]. Institute on the Environment, University of Minnesota, St. Paul, MN 55108, USA. Random Forests for Global and Regional Agricultural Production Forecasts. Due to its great accuracy, the obtained outputs demonstrate that RF is a useful and unique machine-learning method for agricultural yield projections at both the regional and global levels. Support vector regression and k-nearest neighbour are used in the implementation of the paper (SVG).

3. 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. Hence, information on these meteorological variables was acquired on a monthly basis. Gathering 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 Agricultural Yield Forecasts 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 anticipate upcoming events Machine Learning Techniques Supervised machine learning algorithms can use labelled examples to apply what they have previously learnt to fresh 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 unlabelled data to describe a hidden structure. The system does not determine the appropriate output for describing hidden structures from unlabelled data, but instead evaluates the data. The most well-known and effective supervised machine learning algorithm, known as random forest, can perform both classification and regression tasks. It appears to be working 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.

4. IMPLEMENTATION

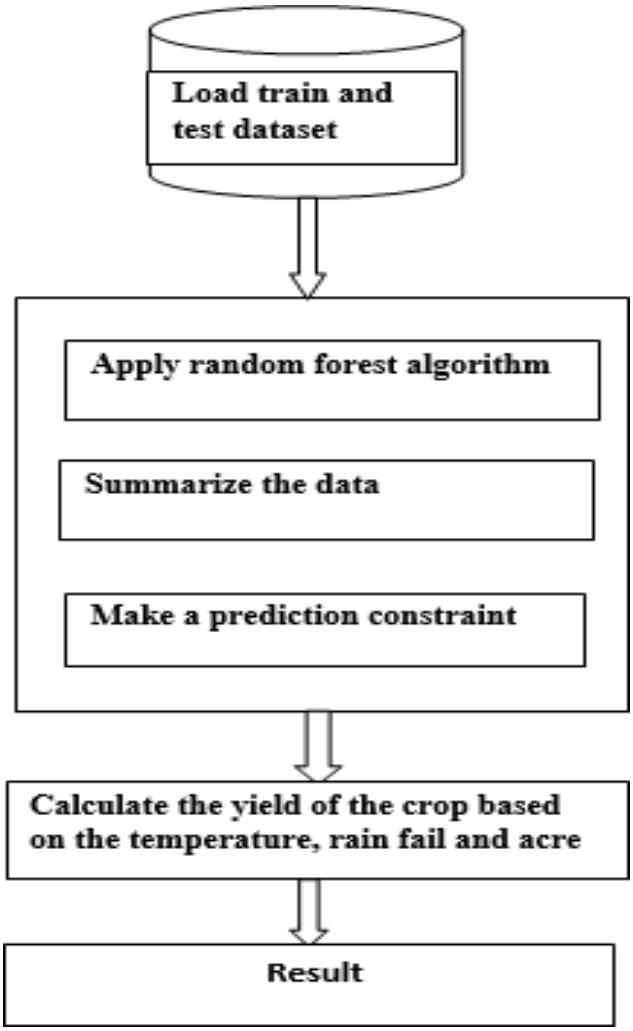


FIGURE 1. Proposed Approach

Fig. 1. Shows the proposed approach and howthe data is summarized, and Random Forest algorithm is applied, and the result is calculated

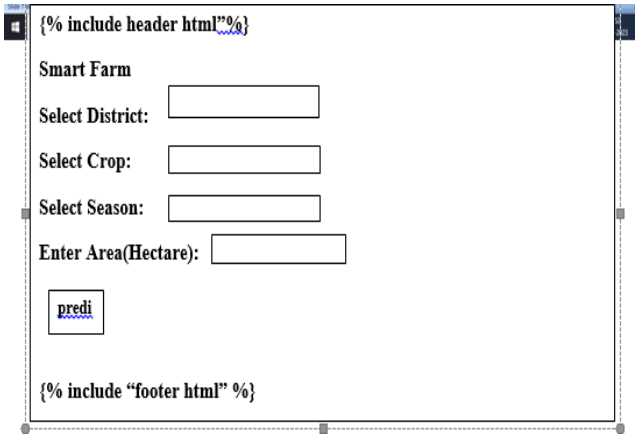


FIGURE 2. Home page

Fig. 2 shows the home pageof 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.

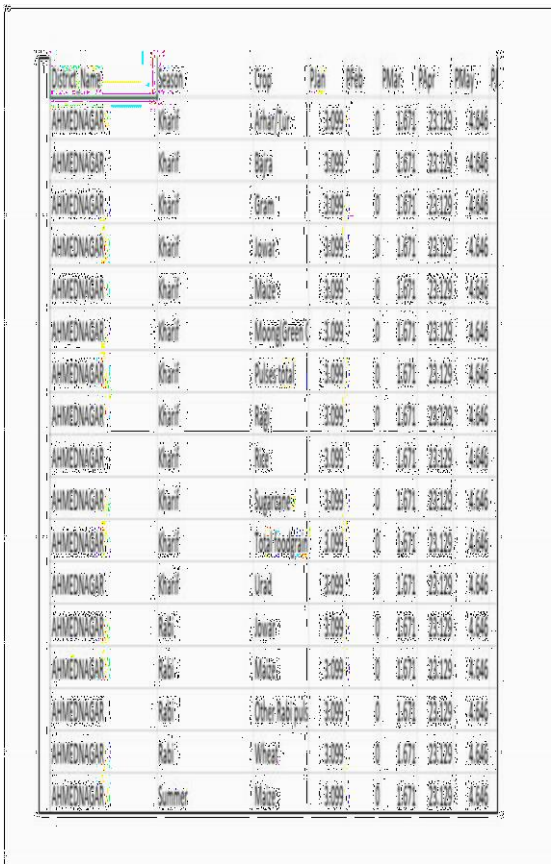


FIGURE 3. Data set

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

5. CONCLUSION AND FUTURE SCOPE

The current study demonstrated the possible use of data mining approaches in agricultural production prediction based on the meteorological input factors. The websitethat was created is user-friendly, and all of thecrops and study districts that were chosen hadpredictions that were more than 75% accurate. Each user can utilise the user- friendly website designed for crop yield prediction by submitting climate data for theirpreferred crop.

REFERENCES

[1]. P.Priya, U.Muthaiah M.Balamurugan.Predicting yield of the crop using machine learning algorithm. International Journal of Engineering Science Research Technology.

[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. GovindarajanMuthukumarasamy.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).

[5]. Shweta K Shahane , Prajakta V Tawale.Prediction On Crop Cultivation. IInternational Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE) Volume 5, Issue 10, October 2016.

[6]. D Ramesh ,B Vishnu Vardhan. Analysis Of Crop Yield Prediction Using Data Mining Techniques. IJRET: International Journal of Research in Engineering and Technology.

[7]. Subhadra Mishra,Debahuti Mishra, Gour Hari Santra. Applications of Machine Learning Techniques in Agricultural Crop Production. Indian Journal of Science and Technology, Vol 9(38), DOI:10.17485/ijst/2016/v9i38/95032, October 2016.

[8]. Konstantinos G. Liakos,Patrizia Busato,Dimitrios Moshou, Simon Pearson ID,Dionysis Bochtis. Machine Learning in Agriculture. Lincoln Institute for Agri-food Technology (LIAT), University of Lincoln, Brayford Way, Brayford Pool,Lincoln LN6 7TS, UK, spearson@lincoln.ac.uk.

[9]. Baisali Ghosh. A Study to Determine Yield for Crop Insurance using Precision Agriculture on an Aerial Platform. Symbiosis Institute of Geoinformatics Symbiosis International University 5th & 6th Floor, Atur Centre, Gokhale Cross Road, Model Colony, Pune – 411016.

[10]. Jig Han Jeong, Jonathan P. Resop, Nathaniel

[11]. D. Mueller, David H. Fleisher ,Kyungdahm Yun, Ethan E. Butler,Soo-Hyung Kim. Random Forests for Global and Regional Crop Yield Predictions. Institute on the Environment, University of Minnesota, St. Paul, MN 55108, United States of America.

[12]. Ecochem Online. (2009). Soil Health and Crop yields. Last modified January 28th 2009.

[13]. Retrieved on March 4th 2009 from http://ecochem. com/healthy_soil. html Food and Agricultural Organization. (2006). The state of Agricultural Commodity Markets. 3739.

[14]. Aditya Shastry, H. A Sanjay And E.

[15]. Bhanushree, “Prediction of crop yield using Regression Technique”, International Journal of computing r12 (2):96- 102 2017, ISSN:1816914]

[16]. E. 14]E. Manjula, S.

[17]. Djodiltachoumy, “A Model for Prediction of Crop Yield”, International Journal of Computational Intelligence and Informatics, Vol. 6: No. 4, March 2017

[18]. Mrs. K. R. Sri Preethaa, S. Nishanthini, D.

[19]. SanthiyaK. Vani Shree, “CropYield Prediction”, International Journal On Engineering Technology and Sciences – IJETS™ISSN(P):2349-3968, ISSN (O):23493976 Volume III, Issue III, March- 2016

[20]. Jharna Majumdar, Sneha Naraseeyappa and Shilpa Ankalaki, “Analysis of agriculture data using data mining techniques: application of big data” Majumdar et al. J Big Data (2017) 4:20 DOI 10. 1186/s40537-017- 0077-4 [16]D. Ramesh and B. Vardhan, “Analysis of crop yield prediction using data mining techniques”, International Journal of Research in Engineering and Technology, vol. 4, no. 1, pp. 47-473, 2015.

[21]. Yethiraj N G, ” Applying data mining techniques in the field of Agriculture and allied sciences”, Vol 01, Issue 02, December 2012. [18]Zelu Zia (2009). An Expert System Based on Spatial Data Mining used Decision Tree for AgricultureLand Grading. Second International Conference on Intelligent Computation Technology and Automation. Oct10-11, China