CROP YIELD AND FERTILIZER PREDICTION USING MACHINE LEARNING

Prof.B.B. Vikhe
Department of Computer Engineering
Pravara Rural Engineering College, Loni
Ahmednagar, India

Pratik Kurhe
Department of Computer Engineering
Pravara Rural Engineering College, Loni
Ahmednagar, India
Pratik.kurhe02@gmail.com

Hrushikesh Magar
Department of Computer Engineering
Pravara Rural Engineering College, Loni
Ahmednagar, India
magarhrushikesh2@gmail.com

Supriya kadu
Department of Computer Engineering
Pravara Rural Engineering College, Loni
Ahmednagar, India
spkadu17@gmail.com

Roshani Sayyad
Department of Computer Engineering
Pravara Rural Engineering College, Loni
Ahmednagar, India
roshnisayyad786@gmail.com

ABSTRACT

In India, agriculture is the primary industry for earning a living. 70% of the Indian workforce is employed directly in Agriculture. The most frequent problem faced by young Indian Farmers is choosing the right crop depending on the needs of the land. They are consequently seeing a marked decline in Productivity. Agriculture is now seriously threatened by climate change and other environmental modifications. One essential strategy for handling this issue practically and successfully is machine learning. Use historical information from the weather, soil, rainfall, and crop yields to predict agricultural yields. We used machine learning algorithms to achieve this. Linear Regression, ANN, K Nearest Neighbors, Random Forest, SVM, and other methods were thought of, but the SVM was selected with an accuracy of 86%. In this project, a web application that can estimateboth overall and individual crop yields is developed. Inaddition, it tells the user when fertilizer should be applied.

Keywords - *Agriculture, Machine Learning, crop-prediction,SVM algorithm, Crop yield*

I. INTRODUCTION

India has the second-highest agricultural output in the world. Many variables, such as the temperature, geography, organics, politics, and economy, have an impact on plant production. Growing more than one crop may be difficult for farmers, especially if they are unaware of market prices.

Wikipedia states that in the decade years prior, India's farmer suicide rate ranged between 1.4 and 1.8 suicides per 100,000 people. Farmer suicides increased from 5,650 in 2014 to over 8,000 in 2015. Using technology to promote cultivation consciousness has become a need. Food insecurity can be caused by seasonal climate variations that affect vital resources including land, water, and air.

A sophisticated system is needed to deal with declining crop yields when crop yields cannot sustainably meet demand. To solve this problem, we propose a system that would allow Farmers to choose crops based on economic and environmental Factors in order to maximize productivity and help the nation's rising food demands. The suggested technique makes use of machine learning to provide predictions. The technology will Give farmers the highest yields by basing crop output and Selection on the appropriate climatic conditions for the crop. The algorithm makes predictions about crop yields based on Variables like rainfall, temperature, area (in hectares), season, and so forth. The method aids in figuring out when fertilizer Should be applied. Forecasting crop yields is a major problem In agriculture.

Every farmer is continually wondering if the productionWill meet their expectations. A farmer's prior knowledge of a Particular crop was previously considered when estimatingYields. The weather, pests, and harvesting schedules all have a Significant impact on agricultural productivity. For riskManagement decisions, accurate data on past agriculturalyields is essential. Farming management. The creation of theseElements and adherence to the relevant standards will be the Responsibility of the trainers.

II. OBJECTIVES

The goal of this study is to forecast agricultural production Under specific weather scenarios and, as a result, suggest suitable crops for that land. It consists of the following steps.

1. Before cleaning the data, gather meteorological data, crop Yield data, soil type data, and rainfall data and combine them In an organized way. By removing erroneous, incomplete, and Nonsensical data, data cleaning increases data quality and, by Extension, productivity.
2. Perform exploratory data analysis (EDA), which aids in Examining the full dataset and summarizing the essential Characteristics. It is used to find trends, spot anomalies, and Create graphical representations of different properties. Most Importantly, it provides us with crucial information about the Significance of each characteristic, how it is related to the class Attribute and other pertinent details.
3. Create training and testing sets from the studied crop data, And then train the model on the training set to predict crop Output given inputs.
4. Compare various algorithms by putting the examined Dataset through them, then figuring out each one's error rate And accuracy. Select the method that has the highest accuracy And the lowest mistake rate.
5. Include the algorithm in a system that is designed to be a Mobile application.

III. PROBLEM STATEMENT

The application of machine learning techniques to forecast crop yield is central to the issue statement. The goal of the project is to help users choose the best crops to plant in order to increase yield and profit. The proposed system aims to overcome the limitations of existing systems and produce predictions by analyzing structured data. The solution we propose is to develop a system that better chooses crops that can be grown throughout the season by taking into account the factors that have the greatest influence on crop productivity. In order to choose high-yielding crops, maximize income, and reduce suicide rates, farmers will benefit from this.

There are two main modules that make up the system:

i. Yield Prediction Module: The user has two options in the yield prediction module: he may choose a specific crop and get its yield, or he can view the top five crops with the highest yield overall.
ii. The Fertilizer Module: This module helps the user decide when to use fertilizer and when not to.

IV. LITERATURE REVIEW

In order to provide defined compounds for suitable surrogates, Our technique automatically collects the necessary weather and Nutritional condition data. Their method's ability to cover large Areas and provide resolution predictions compatible with the System's ground event data files with the highest resolution is Another benefit. Prediction of the harvest season's intensity Before it starts. Crop yield forecasting is done using machine Learning techniques.

In [2]. International journal of technical innovation and Scientific research. This essay's objective is to use the random Forest method to predict crop yield based on the data at hand. The model was created using actual data from Tamil Nadu ,and samples were used for testing. It is possible to predict Agricultural yields with reliability using random forestry Techniques... In-depth study on the use of machine learning in agricultural Production systems is examined in this article

[3]. The Development of machine learning (ML), digitalization, Methodology and augmented computing have opened up new Possibilities for identifying, measuring, and evaluating data-Intensive agricultural processes. Support vector machines were Used to create this article (SVp). The research that follows

[4] looks at the benefits of using aerial devices in agriculturalmarketing for efficient farming. 5th & 6th Floor University, Symbiosis International, Artur Center, Gokhale Cross Road, Model Colony, Pune-411016. GIS and remote sensing are used in precision agriculture to identify changes in the field and take appropriate action (PA). Crop growth fluctuations in agricultural nations can be caused by a variety of reasons, including crop stress, irrigation systems, the prevalence of diseases and pests, and others. Using ensemble learning, the document is handled (EL).

V. METHODOLOGY

Python is commonly regarded as an experimental language In the realm of machine learning, hence this method predicts Harvests combining machine learning and Python as a Programming language. Machine learning creates a trained Model based on the input by using historical data and Knowledge to gain expertise. The output is then predicted by The model. The classifier will be more accurate the greater the Dataset. It has been demonstrated that machine learning's Regression and classification techniques perform better than Other statistical models. Agriculture output is completely Influenced by the soil's chemistry, rainfall, terrain, soil type, Temperature and other geographical factors. Enhancing these Traits is essential for raising agricultural output. Which crops Are grown for the most profit are also influenced by market Variables. All factors must be considered in order to forecast

Returns. Because of this, we developed a system that predicts Crop output based on factors like rainfall, temperature, area, Season and so on using machine learning techniques in Agriculture.

Basic Terminology:

i. Dataset: A collection of sample data with characteristicsNecessary for problem solving.
ii. Characteristics: Useful details that help us comprehend the Issue. To aid in learning, this data is fed into a machine learning system.
iii. Model: A phenomenon's representation via an autonomous Learning algorithm (internal model). He learns from the Information given to him throughout training. The algorithmsTraining produced the model.
For instance, a decision tree model will be trained using aDecision tree algorithm.
.

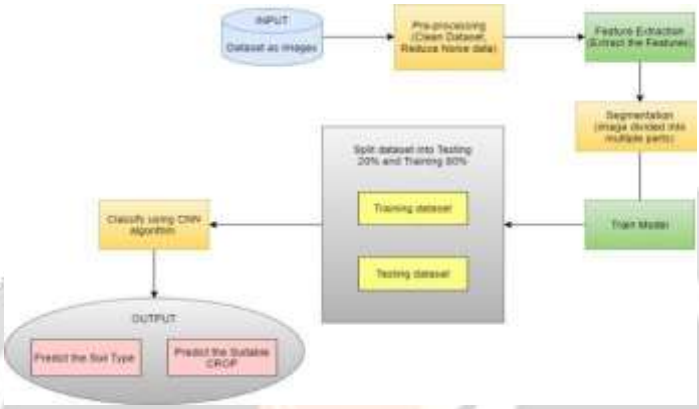

Fig.SystemArchitecture

Basic Process

i. Data collection is the first step in gathering data that the Algorithm can use to make money.
ii. Data preparation comprises identifying important Information, reducing dimensionality, and organizing and Creating data in the best format possible.
iii. Learning: The machine learning algorithm learns by Displaying the data it has gathered and prepared during thisStep, also referred to as the fit phase.
iv. Evaluate: Test the model to see how it performs.

VI. FERTILIZER MODEL

This module's purpose is to provide farmers with guidance on Fertilizer application based on anticipated rainfall over the coming days. To predict rainfall over the next 15 days, we used the "Open Weather" API service. If rain is predicted, farmers are cautioned against fertilizing. The following are thesteps for putting the algorithm into practice.
Step 1: The API contacts the "Open Weather" service when you select this module.
Step 2: From the API call response, read the precipitation overthe next 14 days.
Step 3: Farmers are advised not to fertilize if the amount of Precipitation is greater than 1.25 inches. Otherwise, using Fertilizers is harmless.

VII. DATASETS

Data are crucial to machine learning. The implementation of Algorithm training must include this. By examining historical Data and knowledge, it develops experience. The results will be more precise the more data that is gathered. Gathering data is the first step. This project requires two datasets. Whereas the second is used to forecast fertilizer, the first is utilized to replicate the yield prediction method. These two traits should be used as input information for predicting crop production.Based on weather information from "https://fr.tutiempo.net/" and crop production information from "https://www.kaggle.com/srinivas1/agriculture-crops production-in-India," our dataset. "The yield forecast module dataset needs the following columns because these are the ones that crops depend on the most: State, Year, Crop, Season, Area, Average Temperature, Average Rainfall, Soil Type, Area, and Production. Production" is the dependant or class variable. In this equation, there are 8 independent variables and 1 dependent variable.

State	Year	Season	Crop	Area	Productio	Rainfall	Avg_temp	PH Value	Type of soil
Andaman an	2000	Kharif	Banana	1254	2000	2763.2	16.37	8.5	Sandy soil
Andaman an	2000	Kharif	Banana	2	1	2763.2	16.37	7.5	Sandy soil
Andaman an	2000	Kharif	Rice	102	321	2763.2	16.37	7.5	Sandy soil
Andaman an	2000	Whole Ye	Banana	176	641	2763.2	16.37	7.5	Sandy soil
Andaman an	2000	Whole Ye	Cashewnu	720	165	2763.2	16.37	7.5	Sandy soil
Andaman an	2000	Whole Ye	Coconut	18168	65100000	2763.2	16.37	7.5	Clayey soil
Andaman an	2000	Whole Ye	Dry ginger	36	100	2763.2	15.36	7.5	Loamy soil
Andaman an	2000	Whole Ye	Sugarcane	1	2	2763.2	15.36	7.5	Loamy soil
Andaman an	2000	Whole Ye	Banana	5	15	2763.2	15.36	7.5	Sandy soil
Andaman an	2000	Whole Ye	Banana	40	169	2763.2	15.36	7.5	Sandy soil
Andaman an	2001	Kharif	Banana	1254	2061	3080.9	15.36	7.5	Sandy soil
Andaman an	2001	Kharif	Banana	2	1	3080.9	15.36	7.5	Sandy soil
Andaman an	2001	Kharif	Rice	81	300	3080.9	16.06	7.5	Sandy soil
Andaman an	2001	Whole Ye	Cashewnu	719	192	3080.9	16.06	7.5	Sandy soil
Andaman an	2001	Whole Ye	Coconut	18190	64430000	3080.9	16.06	7.5	Sandy soil
Andaman an	2001	Whole Ye	Dry ginger	46	100	3080.9	16.06	7.5	Sandy soil
Andaman an	2001	Whole Ye	Sugarcane	1	1	3080.9	16.06	7.5	Sandy soil
Andaman an	2001	Whole Ye	Banana	11	33	3080.9	16.06	7.5	Sandy soil
Andaman an	2002	Kharif	Rice	189.2	510.84	2620.2	16.05	7.5	Sandy soil
Andaman an	2002	Whole Ye	Banana	1258	2083	2620.2	16.05	7.5	Sandy soil
Andaman an	2002	Whole Ye	Banana	213	1278	2620.2	16.05	7.5	Sandy soil
Andaman an	2002	Whole Ye	Banana	61	113	2620.2	16.05	7.5	Sandy soil
Maharashtra	2002	Whole Ye	Cashewnu	719	208	2620.2	16.05	7.5	Sandy soil
Maharashtra	2002	Whole Ye	Coconut	18240	67490000	2620.2	16.96	7.5	Sandy soil
Maharashtra	2002	Whole Ye	Dry chilli	413	28.8	2620.2	16.96	7.5	Sandy soil
Maharashtra	2002	Whole Ye	Dry ginger	47.2	133	2620.2	16.96	8.5	Sandy soil

Fig. Merged Dataset

We did this by combining datasets. Location and Dataset Arecombined as a common characteristic for both.

VIII. ALGORITHMS

There are numerous algorithms available for machine learning. Usual techniques include classification, regression, grouping, And association. Whereas clustering and association are Instances of unsupervised learning, classification and regression algorithms are examples of supervised learning. Several categories can be used to group certain algorithms. The most pertinent kind of method will be selected based on the issue description and the intended project outcome. Before choosing and using one, we examined a number of algorithms and assessed their error rates and accuracy.
Search summarization algorithms:
1. SVM (Support Vector Machine) 86%
2. 85% Random Forest

3. An 80 percent convolutional neural network we can infer From the foregoing that the SVM method offers the highest Accuracy for our data set.

IX. CONCLUSION

The idea was devised to help farmers build better economies And to combat the rising rate of farmer suicide. Systems for Recommending crops help farmers choose which crops to Cultivate and anticipate crop output. Also, it advises the user When to apply fertilizer. The right data sets were obtained, Analyzed, and trained using machine learning technology. The System keeps track of the user's position and uses that Information to retrieve the pertinent data from the backend. Users are thus needed to provide fundamental data like area And soil type. This technique, which forecasts the weather 14 days in advance, is valuable in agriculture. The approach also Provides a list of crops based on weather conditions. Further Work will focus on providing crop planting sequences based on soil and weather conditions, as well as regularly updating the data to produce accurate projections. A fully automated system that accomplishes the same task is what the future of work seeks to achieve. Giving the proper fertilisers for a certain crop and region is another item we are working on. Fertilizers and their interactions with soil and climate need to Be studied in this regard. In addition, we want to be ready for Emergencies like the recent rise in onion prices. The system's Ability to give users advice on when to apply fertilizer is one of its most crucial and novel features.

X. REFERENCES

[1]P.Priya, U.Muthaiah M.Balamurugan.Predicting yield ofthe crop using machine learning algorithm. International Journal of Engineering Science Research Technology(IJESRT),2018.
[2]D Ramesh ,B Vishnu Vardhan. Analysis of Crop Yield Prediction Using Data Mining Techniques. IJRET: International Journal of Research in Engineering and Technology. (IJERT), 2015.
[3]] "Indices gallery—ArcGIS Pro | Documentation", Pro.arcgis.com,2020.[Online].Available:https://pro.arcgis.com /en/pro app/help/data/imagery/indices-gallery.html.[Accessed:06- Dec- 2020]
[24] The Land Processes Distributed Active Archive Center(2019).[Online]Available:https://lpdaac.usgs.go v/products/mod09a1v006/

[25] The Land Processes Distributed Active Archive Center(2019). [Online] Available: https://lpdaac.usgs.gov/products/mod11a2v006/
[26] The Land Processes Distributed Active Archive Center (2019). [Online] Available: https://lpdaac.usgs.gov/products/mcd12q1v006//wiki/Agricult ure_in_India
[10]https://www.agricultureinformation.com/forums/threads/a griculture-in-india.101537/
[11]https://en.wikipedia.org/wiki/Farmers%27_suicides_in_In dia
[12] https://towardsdatascience.com/machine-learning-an- introduction-23b84d51e6d0
[13] Groundnut Crop Yield Prediction Using MachineLearning Techniques by Vinita Shah and Prachi Shah.
[14] Developing regression model to forecast the rice yield at Raipur condition by A Jain, JL Chaudhary, MK Beck and Love Kumar
[15] Machine learning approach for forecasting crop yield based on climatic parameters by S.Veenadhari, Dr. Bharat Misra & Dr. CD Singh, International Conference on ComputerCommunication and Informatics (ICCCI -2014), Jan, 2014
[16] Predicting Yield of the Crop Using Machine Learning Algorithm by P.Priya, U.Muthaiah & M.Balamurugan, International Journal of Engineering Sciences & Research International Journal of Engineering Sciences & Research Technology (IJESRT), April, 2018