Crop Recommendation System

1. Abstract

Encouraged to enter generally, agriculture is the backbone of India and also plays a crucial role within the Indian economy by providing a particular percentage of domestic products to make sure food security. But nowadays, food production and prediction are getting depleted thanks to unnatural climatic changes, which can adversely affect the economy of farmers by getting a poor yield and help the farmers stay less familiar in forecasting the long run crops[6]. This research work helps beginner farmers in such a way as to guide them in sowing reasonable crops by deploying machine learning, one of all the advanced technologies in crop prediction. Random Forest, a machine learning algorithm puts forth within the thanks to achieving it[11]. The seed data of the crops are collected here, with suitable parameters like temperature, humidity, and moisture content, which helps the crops to realize a successful growth. additionally, to the software, a web application is being developed. The users' parameters like temperature and their soil condition are going to be taken from the user during this web application so as to begin the prediction process.

Key words:-Machine Learning, Random Forest Algorithm,Django

2. Introduction

Weather plays a vital role in agriculture production. For optimal productivity at a given condition crops must be such their weather requirement matches this weather system. So we want to plan them per the atmospheric condition and soil fertility[9]. this can be a prototype for a crop recommendation algorithm in Python using Machine Learning and Data Analytics. This work presents a system, within the style of an internet site. The business logic in Python uses Machine Learning techniques to predict the foremost profitable crop within the forecasted weather and soil conditions[4]. The proposed system will integrate the info obtained from the soil by applying the machine learning algorithm. This provides a farmer with a spread of options of crops that will be cultivated.

Crop prediction is one of the challenging problems in precision agriculture, and an abundance of models are proposed and validated to date.

This problem requires the utilization of several datasets since crop yield depends on many different factors like climate, weather[9], soil, use of fertilizer, etc. To develop a Crop prediction system [4].

Being a complete software solution, it doesn't allow maintenance factors to be considered much.

Also, the accuracy level of the model will be higher when compared to hardware-based solutions, because the components like soil composition, rainfall value[3], pH value, and weather all get in the picture during the prediction process of the crop.

3. LITERATURE SURVEY

Fatin Farhan Haque, Ahmed Abdelgawad, Venkata Prasanth Yanambaka, and Kumar Yelamarthi stated that agriculture isn't solely an enormous side of the growing economy, however, it's essential for the United States of America to survive in their paper "Crop Yield Analysis mistreatment Machine Learning Algorithms". Predicting crop yield isn't a simple task, because it depends on several parameters like water, ultra-violet (UV), pesticides, fertilizer[1], and therefore the space of the land coated for that region. during this paper, 2 completely different Machine Learning (ML) algorithms area unit projected to investigate the crops' yield. These 2 algorithms, Support Vector Regression (SVR) and regression toward the mean (LR), area unit quite appropriate for validatory the variable parameters within the predicting the continual variable estimation with one hundred forty knowledge points that were non-inheritable. The parameters mentioned higher than area unit key factors poignant the yield of crops. The error rate was measured with the assistance of Mean sq. Error (MSE) and constant of Determination (R2), wherever MSE gave out about zero.005 and R2 gave around zero.85. the constant dataset has been used for fast comparison between the algorithms' performances.[1]

Sachee Nene, Priya, R, and Rushika Ghadge studied the systems to cultivate proper crops for better yield production. They used Kohonen Self Organizing Map (Kohenon's SOM) and BPN (Back Propagation Network) to achieve their objective. This will compare the accuracy obtained from different network learning techniques and then the most accurate one will be provided to the end-user.[4]

Research on Machine learning methods for crop yield prediction and climate change impact assessment in agriculture presented by Andrew Crane Droesch states that the crops are critically dependent on weather for better yield productivity. To prove this he used both extant parametric approaches and over fully-nonparametric neural networks. For his research, he took data from the gridded surface meteorological dataset (METDATA) of Abatzoglou (2013). During his research, he found out that while there is little difference between the models in scenarios in which yields decline less, the SNN projects substantially less-severe impacts in scenarios where yields decline the most across all models, including most of the models in 2070–99 under RCP8.5. For computing the gradients at each step with the full dataset instead, he

used 'minibatch' gradient descent, where small randomized subsets of the data are called 'minibatch'. Another technique used to improve the training of neural networks is 'dropout'. This technique randomly drops some proportion of the parameters at each layer, during each iteration of the network, and updates are computed only for the weights that remain. [6]

Prof. A. V. Deorankar highlight that within the previous couple of decades researchers have had an interest in land mapping and its classification thanks to numerous reasons in their paper “associate Analytical Approach for Soil and Land arrangement mistreatment Image process”. the explanation for a rise within the focus of the analysis community area unit, is the increasing demand for agricultural land and soil health analysis, because the health of the soil, is crucial for the healthy production of crops. Image classification is one such approach for soil and land health analysis. it's a posh method having the results of assorted factors. This paper has projected the study of current research, the issues it addressed, and its prospects. the stress is targeted at the analytical study of assorted advanced and economical classification mechanisms and techniques. Here, it's been tried to check the factors these approaches have addressed to enhance the accuracy of the classification. correct utilization of the number of options of remotely detected knowledge and choosing the most effective appropriate classifier area unit most significant for rising the accuracy of the classification. The data primarily based on classification or Non-parametric classifiers like call tree classifiers or neural networks have gained additional quality for multisource knowledge classification in recent times. However, there's still the scope of any analysis, to scale back uncertainties within the improvement of accuracy of the Image classification mechanisms.[10]

Shilpa Pande, Prem Kumar Ramesh, and Anmol developed a crop detector using a Machine Learning approach. They tried and tasted different Machine Learning algorithms like Support Vector Machine (SVM), Artificial Neural Network (ANN), Random Forest (RF), Multivariate Linear Regression (MLR), and K-Nearest Neighbour (KNN). Among these used algorithms they got the best results with random forest[11].

4. PROPOSED SYSTEM

Module 1: Preprocessing of Data

Here we have taken the raw dataset and then using pandas duplicate values and null values are removed for optimal accuracy of the algorithm.

Here missing data is handled. Missing Data can occur when no information is provided for one or more items or for a whole unit. Missing Data is a very big problem in a real-life scenarios

Here duplicate data is removed using the Pandas library which is available in Machine Learning. Having duplicate data can give rise to Inefficiency and Lack of Productivity. Sometimes it will have inaccurate reporting and less informed decisions.

Module 2: Train Model

Here to recommend a suitable crop we have used Random Forest Algorithm. It could be a popular machine learning algorithm that belongs to the supervised learning technique.

Feature Extraction(Similar features are extracted from the dataset for Training Model).

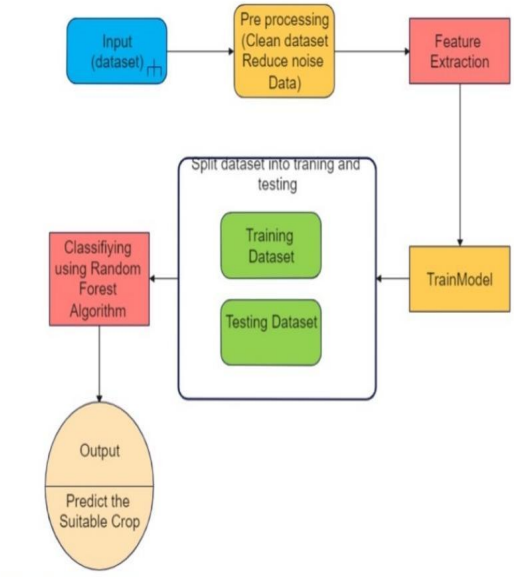
The different features used for the algorithm are Nitrogen, Potassium, Phosphorous, Temperature, Humidity, pH, and Rainfall value. Using these features random forest classifier is trained.

Training Model and splitting it into Training Dataset and Testing Dataset. Now using this training dataset Random Forest could be a classifier that is trained on a variety of decision trees on various subsets of the given dataset and takes the typical to boost the predictive accuracy of that dataset.

Module 3: Suggestion of Crop

-Classifying model using Random Forest Algorithm and Predicting the suitable crop.

We were able to successfully predict the best suitable crop for that particular soil and weather conditions using Random Forest Algorithm. Using this algorithm we got an accuracy of more than 95%.



fig(1): System Architecture

5. PROCEDURAL EXPLANATION

- We initially need to install libraries that are available in python namely pandas, train_test_split, and RandomForestClassifier and they can be downloaded using pip install “library name” in command prompt or python idle or a virtual environment.

- After downloading these libraries we need to import the libraries in our

code by writing import “library name”. If we want to use a particular feature from that library then we need to use the command , from “library_name” import “feature”. If we want to give a name to the library then we can use, import “library_name” as “given_name”. Hence, these libraries would be ready to use in the further code base which we will write.

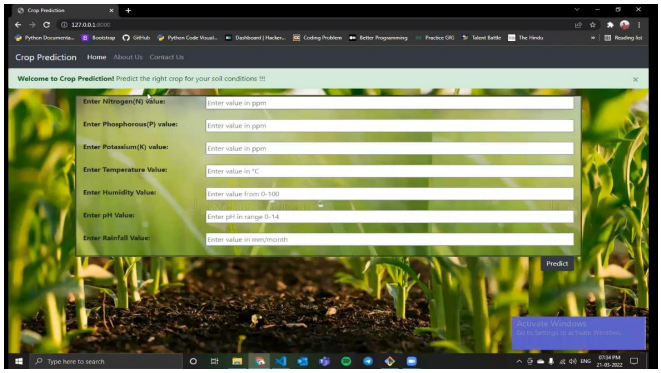
- Further, you can create or download a dataset consisting of 7 features mentioned above. Then read that dataset using pandas.

- In the next step of preprocessing of data, remove the duplicate values and handle the null values.

- Use feature extraction for giving features to the algorithm. Split the dataset into training and testing datasets. Train the classifier using the training dataset.

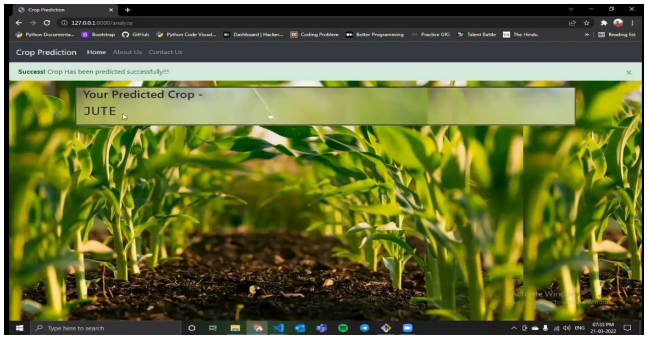
- After the algorithm is trained, to run the Django application go to the current folder, open the command prompt there, and use the command: python manage.py runserver

- Once the server is live enter values of the different attributes. Give those values to the Random Forest algorithm as features for prediction purposes.



fig(2): Taking Input from User

- Now click on predict button to see the best suitable crop for your soil conditions.



fig(3): Prediction of Crop

6.PREVIOUSLY TESTED ALGORITHMS

On conducting several tests on numerous objects our system was able to achieve an

accuracy of 99% which is higher than any other object measurement and detection model previously developed.

PAPER	ALGORITHM	ACCURACY
Efficient Crop Yield Prediction Using Machine Learning Algorithms	Support Vector Machine	75%
Rainfall Prediction using Machine Learning	Artificial Neural Network(ANN)	86%
Prediction of Crop yield using Machine Learning	Support Vector Machine	92%
Machine learning methods for crop yield prediction and climate change impact assessment in agriculture	semiparametric neural networks (SNN)	93.3%
Crop Recommender System Using Machine Learning Approach	Random Forest	95%

7. CONCLUSION AND FUTURE SCOPE

A model is proposed for predicting soil series and providing suitable crop yield suggestions for that specific soil and weather. The model has been tested by applying different styles of algorithms. Random Forest shows highest accuracy in soil classification and suggests crops with less time. It gives us more accuracy as compared to existing systems and offers more benefit to farmers.

In relation to rainfall can depict whether extra water availability is required or not. This research work is enhanced to the next level by availing it to the entire India.

Our future goal is to increase and improve data set with more attributes.

8. References

[1] Arun Kumar, Naveen Kumar, Vishal Vats, “Efficient Crop Yield Prediction Using Machine Learning Algorithms”, International Research Journal of Engineering and Technology (IRJET)-e-ISSN: 2395-0056, pISSN:2395-0072, Volume: 05 Issue: 06— June-2018

[2] Nithin Singh saurabh chaturvedi, “Weather Forecasting Using Machine Learning”, 2019 International Conference on Signal Processing and Communication (ICSC) Volume: 05 — DEC-2019.

[3] Aakash Parmar Mithila Sompura, ”Rainfall Prediction using Machine Learning”, 2017 International Conference on (ICIIECS) at Coimbatore Volume: 3 — March 2017.

[4] Sachee Nene Priya, R “Prediction of Crop yield using Machine Learning”, International Research Journal of Engineering and Technology (IRJET) Volume: 05 Issue: 02 — Feb-2018.

[5] Ramesh Medar Anand M. Ambekar, “Sugarcane Crop prediction Using Supervised Machine Learning” published in International Journal ofIntelligent Systems and Applications Volume: 3 — August 2019.

[6] Andrew Crane Droesch, “Machine learning methods for crop yield prediction and climate change impact assessment in agriculture”, Published by IOP Publishing Ltd Volume: 05 — OCT -2018.

[7] Vinita Shah Prachi Shah, ”Groundnut Prediction Using Machine Learning Techniques “,published in IJSRCSEIT. UGC Journal No : 64718 — March-2020.

[8] Renuka Sujata Terdal, ”Evaluation of Machine Learning Algorithms for Crop Prediction” Published in International Journal of Engineering and Advanced Technology (IJEAT) Volume-8 — August 2019.

[9] P. Vinciya, Dr. A. Valarmathi, “Agriculture Analysis for Next Generation High. Tech Farming in Data Mining” IJARCSSE,vol. 6, Issue 5, 2016

[10] Shivnath Ghosh,Santanu Koley, “Machine Learning for Soil Fertility and Plant Nutrient Management using Back Propagation Neural Networks” IJRITCC, vol. 2, Issue 2,292-297,2014.

[11] Shilpa Pande, Pream Kumar Ramesh, Anmol, “Crop Recommender System Using Machine Learning Approach” IEEE Explore May 2021