

21.10.21_1.docx

by

Submission date: 22-Oct-2021 09:50AM (UTC+0530)

Submission ID: 1680762275

File name: 21.10.21_1.docx (1.12M)

Word count: 6242

Character count: 32273

TITLE - “DigiFarm - A web based one stop solution for farmers”

ABSTRACT

The world population is increasing everyday and in order to keep up with the increasing population an extensive food production system has to be designed. Farming is one of the most important occupations in this modern time. Traditional farming is used by most of the farmers these days, however it has its own disadvantages like less yield leading to lesser revenue. It also makes use of various pesticides which degrades the quality of the soil as well the crops. To solve this problem farmers are nowadays shifting to digital farming which yields more crops thus increasing their revenue. This paper intends to delineate the details of the newly built platform DigiFarm which aids the farmers in making informed decisions based on their soil pattern and weather condition. The central aim of DigiFarm is to help the farmers in predicting the crops most suited for their farm for maximum yield. The crop prediction system is designed in such a way that it is user friendly and it gives accurate results. The prediction model is the result of testing the dataset with several algorithms and methodologies to determine which is best and has the highest accuracy rate. The crops are the target label in the research work and different algorithms like Gradient Boosting Classifier and Random Forest classifier have been used for training the machine learning model and it is constructed carefully using machine learning and artificial intelligence technologies. On testing the crop prediction system which was trained on the dataset, Random Forest model gives nearly hundred percent accuracy. Hence, the Random Forest model is chosen to predict the crops on the website. For deploying this ML model, IBM cloud services have been used. DigiFarm would result in digitization of agriculture on a global scale. However, it would prove extremely beneficial to farmers in India who find it difficult to access the information as their outreach is confined which results in using obsolete methods despite the advancements.

Keywords- Digital Farming, DigiFarm, Farmers, Machine Learning Model, Crop Prediction.

1. INTRODUCTION

Farmers, the soul of the Indian economy, are the ones who feed the huge population of India. Without them, no one would be getting food on their plate today. With a continuous increase in the population, there is a need to increase the yield of the soil to meet the demands of the

twenty-first century world.¹¹ The authors [1] argue that there is a massive need to upgrade the production to meet the world population which is estimated to reach 9.5 billion by the end of 2050. They further note that 80% of the cultivable land grows single crops like rice, corn, wheat and soybean that deteriorate the quality of agriculture across the world.¹¹ [2] Thus, there is a need to find alternatives where agriculture does not suffer because of the depletion and under-utilization of resources. While these literatures summarize the current issues which the agricultural sector is encountering, it does not give innovative details to resolve the same. In other words, the arena of digital farming remains untouched in the traditional purview.

The question arises- how far would digitalization be accepted by the indigenous farmers? According to the author P K Sofia [3] 70% of the population in India depends on traditional farming methods for their livelihood because it is economical and environment friendly to some certain extent. Almost 90% of the tribal population in India is dependent on and acts as guardians of conservative methods of farming [4] [5]. There are several reasons behind the continuation of these methods. For example, according to Lincoln NK [6], traditional farming requires local resources that replenishes the soil constituents and keeps them intact. Scholars like Timothy Johns [7] stated that traditional farming provides financial support to the farmers that secures the diet and nutrition of the global population. The author [8] mentioned that various traditional methods of agriculture have successfully met the overseas demands as well as the domestic demands of crops like rice.

Despite its advantages, traditional farming is not the best method of farming and there are numerous drawbacks for the same. The foremost problem is soil erosion caused due to water, wind, ice or gravity. Most of the above-mentioned papers do not take into consideration that to increase the total yield on a tract of land, traditional farming resorts to the use of excessive pesticides. While approximately a couple of decades ago, the pesticides and fertilizers were hundred percent organic, the farmers increasingly shifted to using synthetic products which depleted the soil quality and was harmful for the crops.

Climatic changes leading to water depletion, air and soil pollution, eutrophication has also created hurdles in the traditional farming methods. Climate change has become one of the most pressing concerns in modern times. The world has indeed experienced some drastic changes in the recent past, from glaciers melting to erratic rainfall. These challenges will continue to grow with the ever-increasing population. In addition to affecting crop growth rate, climate change also affects rainfall variability. They all play a role in raising global temperatures. Figure 1 shows the yearly increase in global temperature from 1880-2020. Global warming has led to a drastic increase in global temperature and it continues to increase every year.

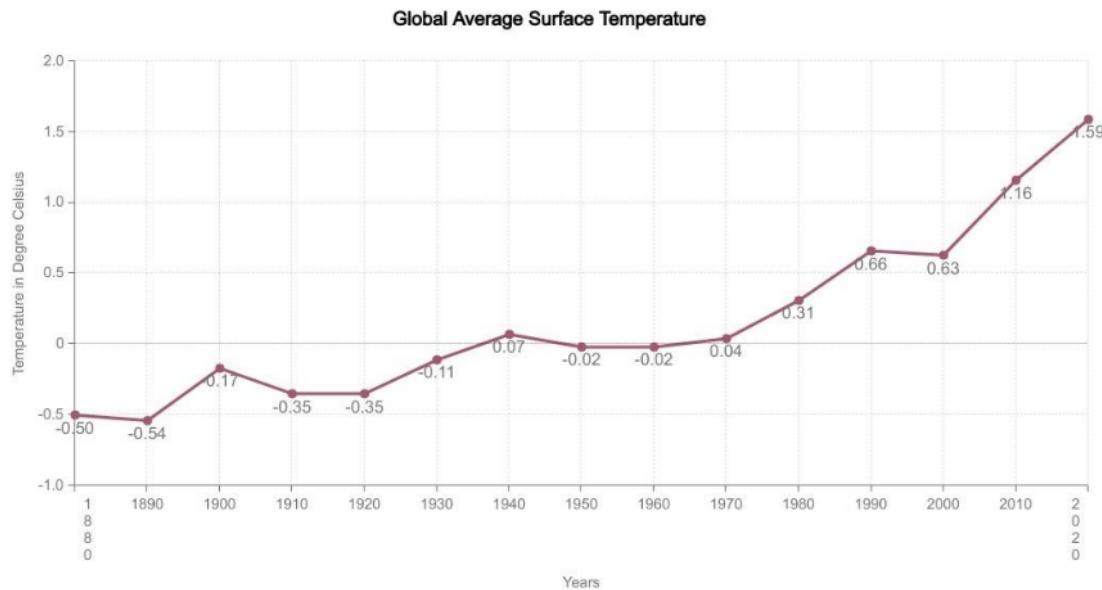


Figure 1: Yearly increase in Global temperature

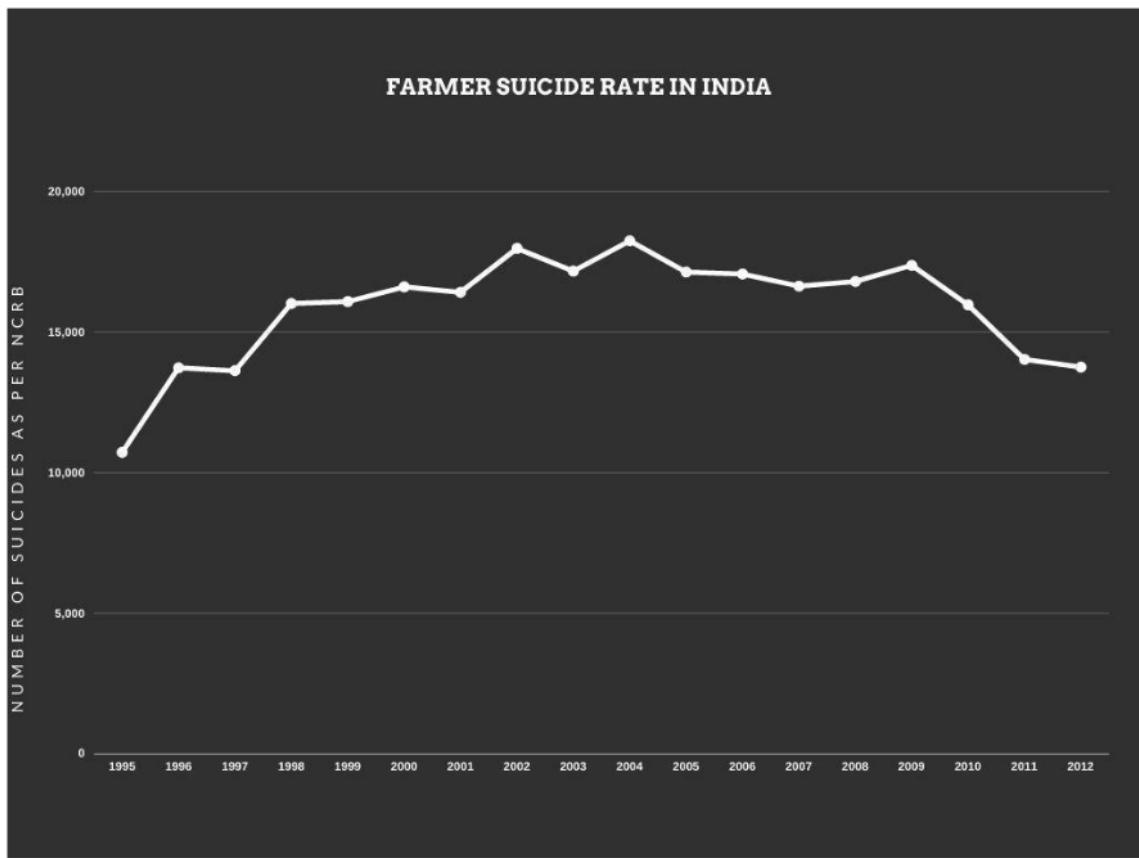


Figure 2: Farmer suicide rate in India as per NCRB report

The changing climate, erratic rainfall, pollution, drawbacks of traditional farming are the central cause behind the magnifying farmer suicide rates all the world every year [9]. Between 1995 and 2006, 166304 farmers committed suicide in India and the number has only increased since then [10]. In order to earn more in a shorter span of time, the farmers start growing cash crops like sugarcane, cotton on the soil that deteriorates its quality. Moreover, the farmers are able to retain a very small percentage of earnings. The cash crops are labour intensive and require more care. Inflation has further added to the troubles of the farmers [11]. Figure 2 shows the farmer suicide rate of India reported as per the National Crime Records Bureau (NCRB) India [12]. This shows the poor conditions of farmers in India and how they have to suffer for revenue and to pay off their debts.

The need of the hour is to transform the way in which farming is viewed. Every sector has not only introduced digitalization but is turning entirely virtual. This is the new world and the

future. In order to keep up with the developments, our primary sector also needs a major revamp which shall be attained with digitalization.

This is where DigiFarm steps in. It is introduced with the sole purpose of helping farmers. The aim of DigiFarm platform is to aid the farmers in selecting the crop that can give them best yield and helping farmers in other aspects. The changes that are caused on their land and soil due to the climatic changes requires an innovative approach. The authors who have previously talked about the latest advancements in agriculture and digital farming such as field scouting and harvesting. They even mention of various concepts such as sensors and robots that can be used extensively in digital farming [13]. Their research is very beneficial for bringing a major change to the current farming trends that can be insightful for the farmers. However, they fail to predict the crops which the farmers could grow for better yield. DigiFarm is primarily aimed at farmers that can assist with their needs. The motive is to make the farmers aware of digital farming and its benefits. By tracking the latest climatic changes, soil patterns and weather conditions, using DigiFarm farmers can predict which crops they can sow on their land for maximum yield.

The prediction is based on the machine learning model which is trained with adequate amounts of statistical data so that the prediction is accurate and is ultimately helpful for the farmers to increase ²² their income. In the literature [14] it is mentioned that the authors used the algorithm that utilizes the arrangement calculation of Help Vector Machine to improve ²² the crop prediction module. The algorithm detects the dirt and gives information about which crop would be best suited and it also tells which nutrients are not present abundantly for that specific crop. Authors Dhivya Elavarasan and P.M. Durairaj Vincent [15] have created a Deep learning based crop yield prediction model and authors [16] have created a ML model for crop yield prediction. In both the cases [15] and [16] they have emphasised on crop yield prediction. Authors [17] have made use of ML ¹³ algorithms Support Vector Machines (SVM), k-Nearest Neighbour to prepare a soil classification and crop prediction model. Researchers [18] have created a crop prediction model by making use of Artificial Neural Network approach where one can predict the crop using parameters such as rainfall, humidity, copper, calcium, magnesium etc.,

The main aim of DigiFarm is to help the farmers in growing their business by growing the right crop at the right time and making them aware about Digital Farming. Due to the global accessibility of DigiFarm, even a farmer in some remote village can access it and make profit thereby. All their queries will be answered by the chatbot present in the DigiFarm platform. DigiFarm which is designed carefully and accurately to predict the most suitable crop that the farmer can produce in his region. For achieving this we have made use of Machine Learning technologies to develop a crop prediction model using Random Forest and Gradient

Boosting classifier algorithms and the dataset which we have used for training the crop prediction model was procured from the department Indian Chamber of Food and Agriculture (ICFA). In the literature [19] the author says that with the use of machine learning models agriculture can be highly benefited. By the use of artificial intelligence appropriate recommendations are provided about crops. The platform is designed to equip the farmers with digitized farming so that with the aid of this, they can receive precise information about which crops would be most suitable for their land. DigiFarm also provides the latest news related to agriculture in the news section.

18

This paper is organized as follows:

Section 2 presents the Methodology used to develop the DigiFarm platform, collect the data and use different algorithms. Section 3 presents the result and description of the DigiFarm. Section 4 summarizes the conclusion to our study derived from the DigiFarm. Section 5 presents the future scope of the whole DigiFarm platform and how it will benefit the mass.

2. METHODOLOGY

Data is a very important part of any Machine Learning Model. Hence, DigiFarm is a user-friendly website designed in such ⁶way that anyone can use it to predict the best crop that can be grown on their soil. In this paper Machine Learning (ML) and Artificial Intelligence (AI) technologies have been used. The prediction model is the result of testing the dataset with the best ML algorithms Random forest (RF) classifier and Gradient Boosting (GB) classifier algorithms.

Following are the steps that we have followed to create the crop prediction model:

- 1) Data collection
- 2) Data preparation and analysis
- 3) Choosing the algorithm for training the dataset
- 4) Testing the Machine Learning (ML) model and evaluation
- 5) Deploying

1) Data collection

To train this prediction model the ¹⁷dataset has been procured from Kaggle website [20] which is shown in Figure 3. This dataset was built by augmenting datasets of rainfall, climate and fertilizer data available for India which was gathered over the period by Indian Chamber of Food and Agriculture (ICFA), India.

| | A | B | C | D | E | F | G | H |
|----|-----|----|----|-------------|-------------|-------------|-------------|--------|
| 1 | N | P | K | temperature | humidity | ph | rainfall | label |
| 2 | 90 | 42 | 43 | 20.87974371 | 82.00274423 | 6.502985292 | 202.9355362 | rice |
| 3 | 85 | 58 | 41 | 21.77046169 | 80.31964408 | 7.038096361 | 226.6555374 | rice |
| 4 | 83 | 95 | 50 | 26.51682337 | 77.79913575 | 5.50947065 | 108.8547508 | banana |
| 5 | 119 | 90 | 48 | 28.66725136 | 79.59242542 | 5.986442306 | 118.2583441 | banana |
| 6 | 78 | 42 | 42 | 20.13017482 | 81.60487287 | 7.628472891 | 262.7173405 | rice |
| 7 | 69 | 37 | 42 | 23.05804872 | 83.37011772 | 7.073453503 | 251.0549998 | rice |
| 8 | 38 | 15 | 30 | 28.91862016 | 48.13974548 | 5.075504537 | 97.01331604 | mango |
| 9 | 12 | 37 | 30 | 31.09779147 | 47.41196659 | 4.546466109 | 90.28624348 | mango |
| 10 | 38 | 19 | 31 | 34.73823882 | 49.08864345 | 5.855119268 | 90.65022183 | mango |
| 11 | 8 | 33 | 29 | 29.98080499 | 49.48613279 | 6.442393461 | 91.82271568 | mango |

Figure 3: Dataset used to train Machine Learning Model

2) Data preparation and analysis

The Dataset which is shown in Figure 3 has 2201 samples among which 90% (i.e. 1980 samples) have been used for the purpose of training. The remaining 10% (i.e. 221 samples) are used for testing purposes. The dataset contains 8 attributes (they are Nitrogen, Phosphorous, Potassium,

Temperature and humidity of the region, pH of the soil, Rainfall in mm in that region and the crop name. The dataset contains 22 distinct categories (i.e., Apple, Banana, Blackgram, Chickpea, Coffee, Cotton, Grapes, Jute, Kidneybeans, Lentil, Maize, Mango, Mothbeans, Mungbean, Muskmelon, Orange, Papaya, Pomegranate, Watermelon, Rice and Coconut).

Before training the dataset the following preliminary analysis of the dataset was conducted:

- i) Finding out some statistical information about the data which is summarized in the Table 1

| Statistical Parameters | Nitrogen | Phosphorus | Potassium | Temperature | Humidity | ph | Rainfall |
|-------------------------------|-----------------|-------------------|------------------|--------------------|-----------------|-----------|-----------------|
| Count | 2200.0 | 2200.0 | 2200.0 | 2200.0 | 2200.0 | 2200.0 | 2200.0 |
| Mean | 50.55 | 53.36 | 48.14 | 25.61 | 71.48 | 6.48 | 103.46 |
| Standard Deviation | 36.91 | 32.98 | 50.64 | 5.063 | 22.26 | 0.77 | 54.95 |
| Minimum Value | 0.0 | 5.0 | 5.0 | 8.82 | 14.25 | 3.50 | 20.21 |
| 25% | 21.0 | 28.0 | 20.0 | 22.76 | 60.26 | 5.97 | 64.55 |
| 50% | 37.0 | 51.0 | 32.0 | 25.59 | 80.47 | 6.42 | 94.86 |
| 75% | 84.25 | 68.00 | 49.0 | 28.56 | 89.94 | 6.92 | 124.26 |
| Maximum Value | 140.0 | 145.0 | 205.0 | 43.67 | 99.98 | 9.93 | 298.56 |

Table 1: Basic statistical information about the dataset

ii) The correlation between different attributes which is summarized in the following Table 2 was found.

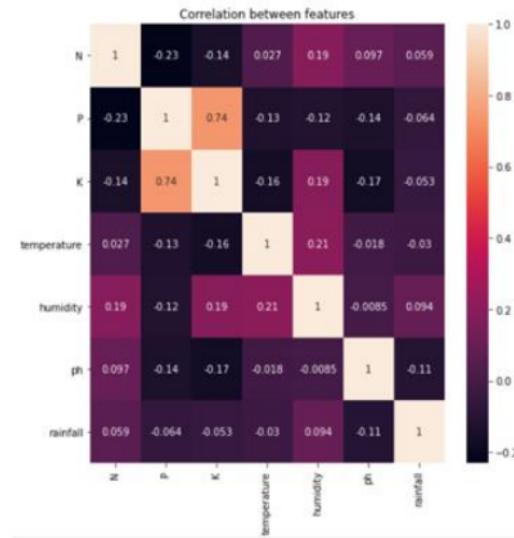


Table 2: Correlation between features

iii) Since there were no missing cells or null values to the next step i.e., choosing the algorithm and training it the next step was then carried out.

3) Choosing the algorithm for training the dataset

In this research work, the “crop name” has been considered as a target label and used “Multiclass Classification”. Gradient Boosting Classifier and Random forest algorithms for training the ML model have been used.

19

Gradient Boosting (GB) Classifier Algorithm

Gradient boosting is a Machine Learning method used for classification, regression and other tasks which was invented by Leo Breiman. It is a boosted ensemble of tree as opposed to a bagged ensemble they have very low interpretability because the second tree in the model may or may not predict the same target as the previous model and that the subsequent trees in the model seek to predict how far off the original predictions were from the truth by using the residuals from the prior trees. In this way each subsequent tree in the GB model slowly reduces the overall error of the previous trees. This enables GB models to have very high predictive power but low interpretability. In addition GB models are quite prone to overfitting the training data to combat there are several

extra hyper parameters that are not needed in forests. They are learning rate which controls how you add subsequent trees together and also regularisation in the form of ridge and lasso hyper parameters.

Random Forest (RF) Classifier Algorithm

The Random decision forests or Random forests are an ensemble method (it groups multiple Decision tree predictors) used for regression, classification and other ML tasks which was developed by Leo Breiman. In this each tree makes their own prediction and they are aggregated into a final prediction either by votes for classification problems or as an average for regression problems. If all of the decision trees are the same then each tree will predict the same output; this is where the random part of the random forests come into play. There are two aspects of randomness involved. First is what features in each decision tree a random subset of features is chosen. Second aspect of randomness is using only a sample of the training data each time a tree is fit. The idea is to have each row and feature utilized in at least one of the decision trees, however not to use all the features in any single decision tree. This allowed to build trees that are not correlated while adding variation to our collection of models and reduce the risk of overfitting. Even with this approach overfitting is still a concern that can be solved by setting the max depth. Max depth is the number of questions asked before we reach the prediction. We limit the depth to reduce the risk of overfitting.

4) Testing the Machine Learning (ML) model and evaluation

On training the GB classifier algorithm and RF classifier algorithm models by the dataset, the accuracy from GB algorithm was 0.996 and that of RF was 0.998. On testing both the models with 10% samples of the dataset, it is found that the accuracy from the GB Classifier algorithm is 0.982 and that of the RF classifier was 0.989. Since the accuracy of RF was high in both training and testing phases an RF model was chosen to deploy and use it for predicting the crop.

5) Deploying

For deploying this ML model IBM cloud services have been used. Since it is deployed in a cloud, the ML Model for prediction through Application Program Interface (API) has been used.

3. RESULT AND DISCUSSION

In this section the outcome of the platform “DigiFarm” which is designed to carefully and accurately predict the most suitable crop that the farmer can produce in his region is discussed. DigiFarm is designed to equip the farmers with digitized farming so that they can make the most out of their crops. With the aid of this platform, they can receive precise information about which crops would be most suitable for their land. User can predict their crop based on two methods:

Method 1: By making use of region's weather conditions, pH value of the soil, rainfall pattern and soil composition (i.e., nitrogen, phosphorus and potassium) as inputs

Method 2: By making use of place/location and current season as inputs

3.1 Software Compatibility

The platform DigiFarm is compatible with latest versions of browsers such as Google Chrome, Microsoft Edge, Mozilla Firefox etc. The Front-end of DigiFarm is designed using HTML 5, CSS 3 and JavaScript whereas for the Back-end we have used Django 3.2.5 (i.e., a Python back-end framework).

3.2 Home Page



Figure 4: Home Page of the DigiFarm

DigiFarm platform along with its name is shown in Figure 4. The logo displayed on the top left corner has two central elements: a hand and the water. The water is representative of rivers and oceans that forms the backbone of the irrigation system in Indian agriculture. The hand that is holding a plant represents the farmers who grow the crops. Since India is an agrarian economy, not only is the population dependent on the farmers for food, but the national economy is also dependent on the yield from the primary sector.

This page connects the user to all the different pages on the platform. The Navigation bar on the homepage as shown in Figure 4 has different buttons for various purposes such as ChatBot, Prediction and News. Additionally, there are two buttons at the center of the homepage: one to get the detailed tutorial on how to use the website and make the best use of it for the users and another button is to predict the most suited crop(s). The author [21] stated that the user interface should be simple to grasp for the user to understand it. Hence, DigiFarm platform has a very simple User Interface (UI) because of which users i.e. mainly farmers can access and navigate through our platform to get the best out of it without facing any hurdles.

The following sections have been discussed in detail:

3.3 Prediction Section

The prediction section is the main part of DigiFarm. The “Prediction” section as shown in Figure 5 can be used to predict the most suitable crops that can be sown on their land by using two different methods.

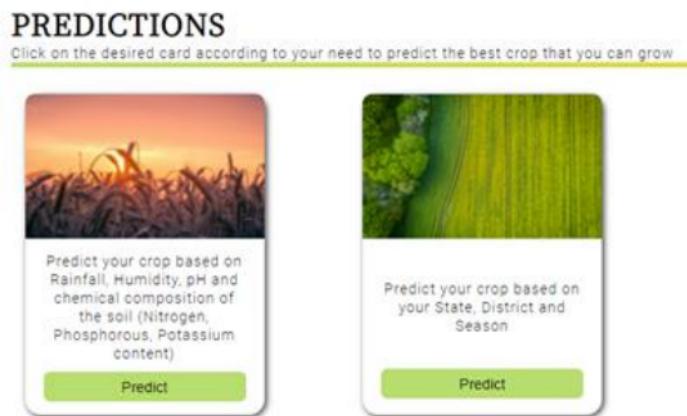


Figure 5: Predictions Section (i.e. Different methods for predicting the crop)

3.3.1 Method 1

As shown in the Figure 5(a) users can predict the crop by entering details such as rainfall, humidity, and temperature, and pH value along with the composition of the soil (i.e. nitrogen, phosphorus, and potassium content), The system will predict the crop which is most suitable for the given geographical conditions.

Instructions

In order to predict the best crop, Enter the values for the following fields. You have to get Nitrogen, Potassium, Phosphorous and pH values of the soil by contacting your nearest Government Soil Testing Laboratory. Incase if you're facing any difficulties please free to open our ChatBot by clicking the button at the bottom of the screen and get your problem solved.

NOTE: Fields marked with (*) are compulsory.

| | |
|----------------------|---|
| Nitrogen Content* | Location* |
| N | Enter your place name. Example: bengaluru |
| Phosphorous Content* | Temperature in °C |
| P | Example: 25 |
| Potassium Content* | Humidity |
| K | Example: 60 |
| pH of the soil* | Submit |
| Example: 7 | |
| Rainfall in mm* | |
| Example: 900 | |

Figure 5(a): Prediction of crops using method 1

Consider an instance, when the user enters the values of nitrogen content as 20, phosphorus content as 89, potassium content as 40, pH value as 6, rainfall as 700, location as Bengaluru the Machine Learning (ML) model will predict the crop based on the values entered and the result is displayed as Coconut as shown in the Figure 5(b).

Result

The best crop that can be sown on your soil is **Coconut**

Figure 5(b): The output based on method 1

Similarly, for if the input values were 86 for nitrogen, 76 for phosphorous, 54 for potassium, 29°C for temperature, 80% humidity, 5.9 for ph and rainfall is 90mm then our model accurately predicted Rice as the best crop that can be sown (as shown in the Figure 5(c)) based on the conditions mentioned.

Result

The best crop that can be sown on your soil is **RICE**

Figure 5(c): Output based on method 1

3.2.1.1 Discussion on prediction of crop using method 1

Gradient Boost Classifier and Random Forest classifier algorithms that have been used for training the dataset are explained in section 2 (i.e. Methodology Section). In ML a confusion matrix is a table that describes the ways in which any ML model is confused while making its predictions. The confusion matrix that was generated on running the test data with Gradient Boost model and Random Forest model is as shown in the Figure 6 and 7 respectively where the numberings 0-21 refers to the crops as shown in the lookup table 3. For instance the Gradient Boosting Model is confused between crop 8 i.e. kidney beans and crop 20 i.e. rice 7 times (in other words this model wrongly predicted the crop 8 as crop 20, 7 times).

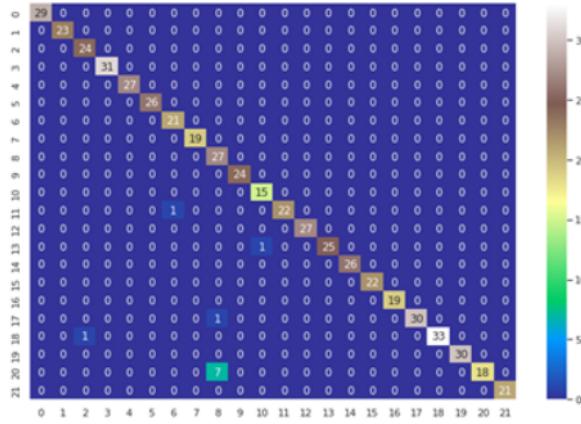


Figure 6: Confusion Matrix for GB Model

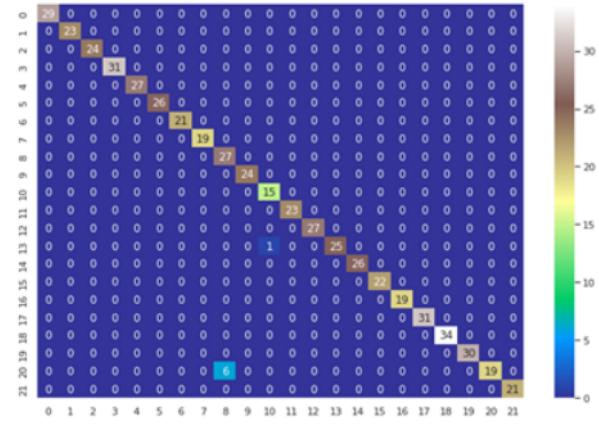


Figure 7: Confusion Matrix for RF Model

| Index | Corresponding Crop | Index | Corresponding Crop | Index | Corresponding Crop |
|-------|--------------------|-------|--------------------|-------|--------------------|
| 0 | Apple | 8 | Kidneybeans | 16 | Papaya |
| 1 | Banana | 9 | Lentil | 17 | Pigeonpea |
| 2 | Blackgram | 10 | Maize | 18 | Pomegranate |
| 3 | Chickpea | 11 | Mango | 19 | Watermelon |

| | | | | | |
|---|--------|----|-----------|----|---------|
| 4 | Coffee | 12 | Mothbeans | 20 | Rice |
| 5 | Cotton | 13 | Mungbean | 21 | Coconut |
| 6 | Grapes | 14 | Muskmelon | | |
| 7 | Jute | 15 | Orange | | |

Table 3: Crop lookup table

20

A true positive prediction means an outcome where the model has correctly predicted that the value belongs to a positive class. A true negative prediction means an outcome where the model has correctly predicted that the value belongs to a negative class. A false positive is an outcome where the model has incorrectly predicted that the value belongs to a positive class. And false negative is an outcome where the model has incorrectly predicted that the value belongs to a negative class.

15

Table 4 summarizes the true positive rate, true negative rate, false positive rate and false negative rate for Random Forest model and Gradient boosting model.

| Confusion Matrix Parameters | Random Forest Model | Gradient Boosting Model |
|-----------------------------|---------------------|-------------------------|
| True positive rate | 0.990 | 0.981 |
| True negative rate | 0.990 | 0.999 |
| False positive rate | 0.0004 | 0.0008 |
| False negative rate | 0.009 | 0.018 |

Table 4: Different type of prediction rates of both RF and GB model

24

The way the ML models are generally evaluated through the parameters accuracy, precision, recall and F1 score.

16

Accuracy is the ratio of total number of correctly predicted data points to the total number of all the data points i.e.,

23

$$\text{Accuracy} = (\text{True Positives} + \text{True Negatives}) / (\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives})$$

Precision is defined as the ratio of True Positives to the sum of True positives and false positives i.e.,

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Recall is defined as the ratio of True Positives to the sum of True positives and false negatives i.e.,

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

F1 score is a simple way to compare two classifiers and is defined as the harmonic mean of recall and precision.

$$\text{F1 score} = \frac{2}{(1/\text{Precision})+1/\text{Recall}})$$

$$= \frac{\text{True Positives}}{\text{True Positives} + ((\text{False Positives} + \text{False Negatives})/2)}$$

Table 5 summarizes the values for these parameters that we got on testing Random Forest and Gradient Boosting models.

| Performance Measures | Random Forest Model | Gradient Boosting Model |
|----------------------|---------------------|-------------------------|
| Accuracy | 0.989 | 0.982 |
| Recall | 0.990 | 0.981 |
| Precision | 0.990 | 0.981 |
| F1 Score | 0.990 | 0.981 |

Table 5: Performance measures (i.e. Accuracy, Recall, Precision, F1 Score) of predictive models

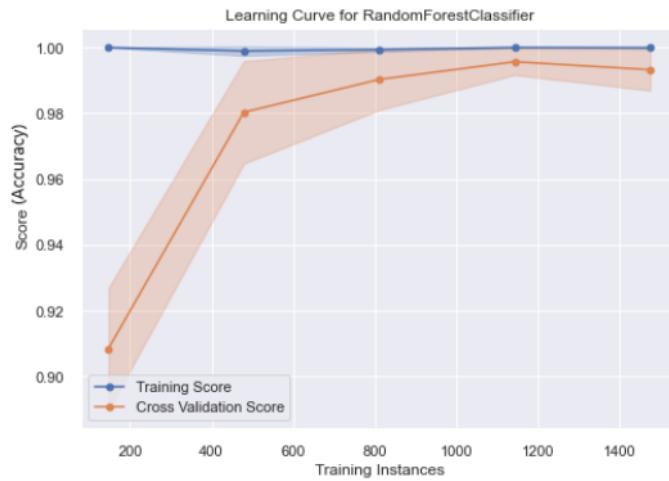


Figure 8: Accuracy curve for RF model for training phase

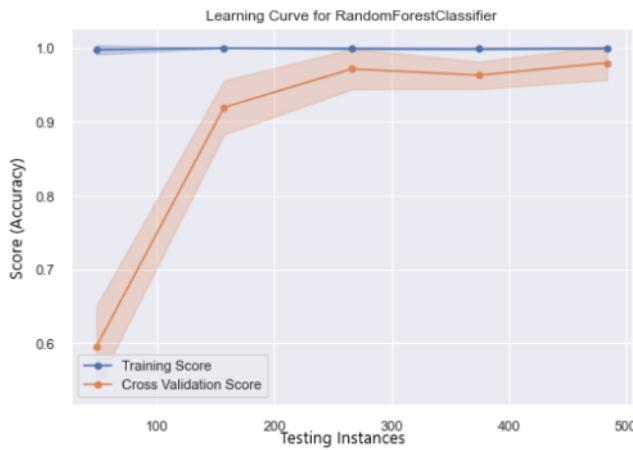


Figure 9: Accuracy curve for RF model for testing phase

Figure 8 and 9 represents the accuracy curve for RF models while training and testing phases respectively.

Figure 10 and 11 represents the accuracy curve for RF model while training and testing phases respectively

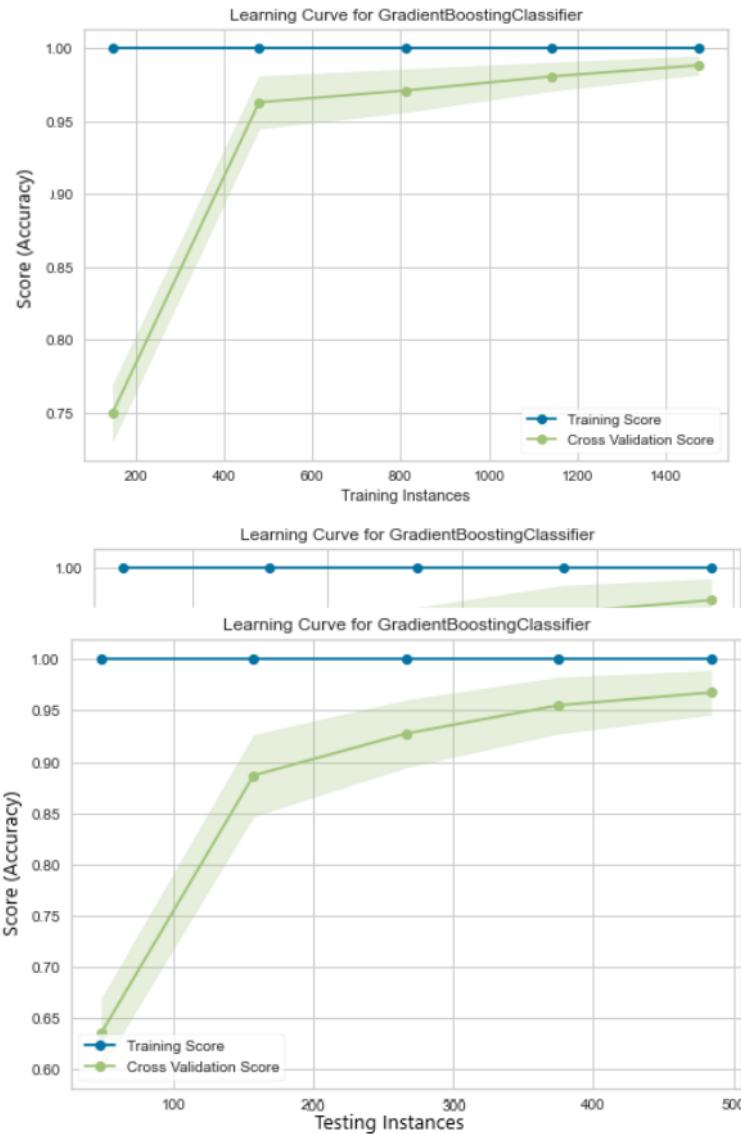


Figure 10: Accuracy curve for GB model for training phase

Figure 11: Accuracy curve for GB model for testing phase

A cross validation is a commonly used method in ML for comparing and selecting different ML predictions models. As it can be seen in the Figure 8 and 9 there is a steep increase in the cross validation score till 500 training instances and till 300 testing instances approximately of RF model, later on the curve tends to saturate. Similarly the cross validation score in the Figure 10 and 11 there is a steep increase in the cross validation score till 500 training instances 300 testing instances approximately for GB model and after which the curve tends to saturate.

Even though the accuracy of RF model and GB model is close to one as shown in Figures 9, 10, 11 and 12, the accuracy (testing accuracy) of RF model is higher than GB model by 0.007 as shown in Table 3.

Authors Disha Chiplonker et al., [22] have created a crop prediction model using machine learning algorithms and techniques such as linear regression and Random Forest regressor. But the accuracy that they have got is 94.63%. whereas the model used in DigiFarm can predict the crop with an accuracy of 98.9%.

As it can be seen in the Table 3 RF Model outperformed Gradient Boosting Model in all other aspects (Recall, Precision and F1 Score) hence RF model is chosen to deploy and use it for predicting the crop. This system of crop prediction results in accuracy and efficiency which is unprecedented.

3.3.2 Method 2

Another option for predicting the crops is by mentioning their state, district, and season as shown in Figure 13(a).

Instructions

In order to predict the best crop, select your State, District and Season from the dropdown box.
NOTE: Fields marked with (*) are compulsory.

State*

Karnataka

District*

BANGALORE RURAL

Season*

Kharif

Submit

Figure 13(a): Input fields for approach 2

For example:- When user enter the state as Karnataka, district as Bengaluru Rural, season as kharif (ref Figure 5(a)), we get the crop prediction of ONION, DRY GINGER, RAGI, BAJRA, MAIZE, RICE, GRAM and many more as shown in Figure 13(b).

Result

The best crops that can be according to the state, district and season specified by you are:

- Niger seed
- Dry ginger
- Groundnut
- Onion
- Rapeseed &Mustard
- Soyabean
- Urad
- Gram
- Cowpea(Lobia)
- Moong(Green Gram)
- Small millets
- Sesamum
- Bajra
- Other Kharif pulses
- Potato
- Paddy
- Ragi
- Castor seed
- Dry chillies
- Rice
- Peas & beans (Pulses)
- Cotton(lint)
- Sunflower
- Horse-gram
- Maize
- Arhar/Tur

Figure 13(b): Sample result from the second method for Crop Prediction

3.3.2.1 Discussion on prediction of crop by method 2

For this second method the dataset which is used is different from the first one. This dataset [23] contains the name of the crops which gave the best yield in each seasons (i.e. Autumn, Kharif, Rabi, Summer, Winter). This crop and season data is available for all the districts of India. This dataset has 4 columns (as shown in Figure 14) namely the State Name, District Name, Cropping Season and crop names.

For predicting the crops in this method ML model is not used but instead the dataset is being searched and the crop names are displayed according to the information (i.e. State name, District name and the preferred season) entered by the user.

| A | B | C | D |
|-------------------------------|---------------|------------|-------------------|
| 1 State_Name | District_Name | Season | Crop |
| 2 Andaman and Nicobar Islands | NICOBARS | Kharif | Arecanut |
| 3 Uttarakhand | PAURI GARHWAL | Rabi | Rapeseed &Mustard |
| 4 Uttarakhand | PAURI GARHWAL | Rabi | Wheat |
| 5 Tamil Nadu | MADURAI | Whole Year | Cashewnut |
| 6 Tamil Nadu | MADURAI | Whole Year | Coconut |
| 7 Tamil Nadu | MADURAI | Whole Year | Coriander |
| 8 Karnataka | KOPPAL | Kharif | Sesamum |
| 9 Karnataka | KOPPAL | Kharif | Small millets |
| 10 Karnataka | KOPPAL | Kharif | Sunflower |

Figure 14: Crop prediction dataset for method 2

Farmers generally wish to continue growing the same crops on their land either to avoid risk or due to lack of awareness. However, DigiFarm would enable them to go beyond their regular pattern by providing accurate and precise information about the crop which they can sow on their land to get the maximum yield.

3.4 AgriBot

¹ In the current growing age of digitization, Artificial Intelligence (AI) powered chatbots are playing a leading role by exemplifying the function of a virtual assistant that could manage a conversation via speech or textual methods [24]. It makes use of voice or textual queries to get answers, perform actions and recommendations according to user needs. As stated ¹ by the author [25] Chatbots are used for a variety of purposes, such as education or assistance. They are adaptable to the user's individual language usages, searches, and preferences with continuing use. With ¹ the advent of AI, virtual assistants can be seen penetrating to the nook and corner of the world [26]. A conversational bot with a voice and/or chat interface can play a principal role in solving ¹ the user queries by giving instant service. This drastically reduces waiting time for the users (i.e. mainly farmers) to get their queries solved. The ChatBot as shown in Figure 4 is an AI based bot called the "AgriBot". It is created using the IBM Watson Chatbot services (i.e. a part of IBM Cloud services).

The result of the AgriBot is the easy accessibility of DigiFarm platform for the users. The AgriBot provides users three options (i.e. as shown in the Figure 15) which can be availed one at a time. This helps the user to get their queries solved.

The different options available for users are:

- 1) To guide the users about the platform and how to make use of it.
- 2) Various methods to predict the crop
- 3) Directs users to a Google Form which they can fill for any further queries

The presence of an AgriBot on this platform simplifies the user experience which is the broader aim of the Digifarm.



Figure 15: AgriBot

3.6 News Segment

In today's time websites have become a major portal for conveying information such as agriculture related news [27]. The news segment results in bringing together the latest information about agriculture from across the globe [28] as shown in Figure 16. The idea behind this is to inform the users about the technological developments happening in the agricultural field from different parts of the world. This enables them to learn from these techniques and apply the suitable ones on their land. The result would be increased productivity, developing a nature of taking risks and also equipping them with the most updated advancements which they may find suitable for their farm.

In the news section of DigiFarm, we have used RSS (Really Simple Syndication) news feed from "The Hindu Agri-Business" section.

NEWS

Find out the latest NEWS about agriculture below

The screenshot shows a news feed from 'The Hindu - Agri-Business'. The first article is titled 'India's agriculture exports can be doubled in 'medium term'', says ITC Chairman' with a timestamp of Thu, 30 Sep 2021 17:49:08 +0530. The second article is 'There is an urgent need to deal with the agriculture value chain, which would help in improving the quality of income in rural areas and ushering in socio-economic transformation in the country,' ITC Chairman Sanjiv Puri said.' with a timestamp of Thu, 30 Sep 2021 17:49:08 +0530. The third article is 'Data : Most farmers sold to private traders in 2019, new survey data shows' with a timestamp of Wed, 15 Sep 2021 10:44:35 +0530. The fourth article is 'Data also show that a majority of the farmers were unaware of the Minimum Support Price regime' with a timestamp of Thu, 09 Sep 2021 08:08:00 +0530. The fifth article is 'Cabinet decision to hike MSP for Rabi crops will ensure max. remunerative price for farmers: PM' with a timestamp of Thu, 09 Sep 2021 08:08:00 +0530. The sixth article is 'Prime Minister Narendra Modi on Wednesday said the Cabinet's approval to increase the Minimum Support Prices (MSP) for Rabi crops will ensure maximum' with a timestamp of Thu, 09 Sep 2021 08:10:27 +0530. The seventh article is 'Amid protests, Centre hikes minimum support price for rabi crops' with a timestamp of Thu, 09 Sep 2021 08:10:27 +0530. At the bottom left is the text 'Courtesy: THE HINDU'. On the right side, there are buttons for 'Feed Widget', 'Activate Windows', and 'Go to Settings to activate Windi'.

Figure 16: News section

The idea is to build a platform which solves all the queries of the farmers holistically and comprehensively. The main aim of DigiFarm is to help the farmers in increasing their production and yield per square by choosing the right crop for their field at the right time. Also, it enables them to sell their crop for maximum revenue. DigiFarm would result in digitization of agriculture on a global scale. However, it would prove extremely beneficial to farmers in India who find it difficult to access the information as their outreach is confined which results in using obsolete methods despite the advancements. DigiFarm, with the aid of technology will help them with the best possible information for their farm using a single platform.

CONCLUSION

Today, farming is considered one of the most important occupations. Traditional farming is considered to be one of the best methods for farming but it has its own drawbacks like less yield of crops leading to less revenue if farmers and also it uses pesticides and harmful fertilizers leading to pollution of crops and soil. Hence in order to avoid these drawbacks digital farming is promoted and it yields more crops thus increasing the revenue of farmers. This paper intends to introduce the new arenas of digital farming. Innovative techniques of farming are required to increase the yield of the tract of land and DigiFarm aims to aid the farmers in this endeavour. By helping the farmers through crop prediction, DigiFarm keeps up with the latest technological advancements. The platform DigiFarm which aids the farmers in making informed decisions based on their soil pattern and weather condition. The central aim of DigiFarm is to help the farmers in predicting the crops most suited for their farm for maximum yield. The crop prediction system is designed in such a way that it is user friendly and it gives accurate results. The prediction model is the result of testing the dataset with several algorithms and methodologies to determine which is best and has the highest

accuracy rate. The prediction is based on the machine learning model which is trained with adequate amounts of statistical data so that the prediction is accurate and is ultimately helpful for the farmers to increase their income. Due to the global accessibility of DigiFarm, even a farmer in some remote village can access it and make profit thereby. The platform is designed to equip the farmers with digitized farming so that with the aid of this, they can receive precise information about which crops would be most suitable for their land. All their queries will be answered by the chatbot present in the DigiFarm platform. DigiFarm is designed carefully and accurately to predict the most suitable crop that the farmer can produce in his region. DigiFarm also provides the latest news related to agriculture in the news section.

5. References

- [1] Patel, Sanoj Kumar, Anil Sharma, and Gopal Shankar Singh. "Traditional agricultural practices in India: an approach for environmental sustainability and food security." *Energy, Ecology and Environment* 5.4 (2020): 253-271.
- [2] FAO (2016) The State of Food and Agriculture, Climate Change, Agriculture and Food Security. Food and Agriculture Organization of the United Nations Rome, 2016. www.fao.org
- [3] Adams MW, Ellingboe AH, Rossman EC (1971) Biological uniformity and disease epidemics. *Bioscience* 21(21):1067–1070
- [4] Sofia, P. K., Rajendra Prasad, and V. K. Vijay. "Organic farming-tradition reinvented." (2006).
- [5] Singh GS, Ram SC, Kuniyal JC (1997) Changing traditional land use patterns in the Great Himalayas: a case study of Lahaul Valley. *J Environ Syst* 25:195–211
- [6] Jeeva SRDN, Laloo RC, Mishra BP (2006) Traditional agricultural practices in Meghalaya, North East India. *Indian J Trad Knowl* 5(1):7–18
- [7] Pradhan, Aliza, et al. "Potential of conservation agriculture (CA) for climate change adaptation and food security under rainfed uplands of India: A transdisciplinary approach." *Agricultural Systems* 163 (2018): 27-35.
- [8] Lincoln NK (2019) Learning from indigenous agriculture. *Nat Sustain* 2(3):167
- [9] Johns, Timothy, Bronwen Powell, Patrick Maundu, and Pablo B. Eyzaguirre. "Agricultural biodiversity as a link between traditional food systems and contemporary development, social integrity and ecological health." *Journal of the Science of Food and Agriculture* 93, no. 14 (2013): 3433-3442.

[10] Rana, Ram Bahadur, Chris Garforth, BhuwonSthapit, and Devra Jarvis. "Influence of socio-economic and cultural factors in rice varietal diversity management on-farm in Nepal." *Agriculture and human values* 24, no. 4 (2007): 461-472.

[11] Nagaraj K. Farmers' suicides in India: magnitudes, trendsand spatial patterns.

[12] Kennedy, Jonathan, and Lawrence King. "The political economy of farmers' suicides in India: indebted cash-crop farmers with marginal landholdings explain state-level variation in suicide rates." *Globalization and health* 10.1 (2014): 1-9.

[13] R Shamshiri, R., Weltzien, C., Hameed, I. A., J Yule, I., E Grift, T., Balasundram, S. K., ... & Chowdhary, G. (2018). Research and development in agricultural robotics: A perspective of digital farming.

[14] Suresh, G., Kumar, A. S., Lekashri, S., & Manikandan, R. (2021). Efficient crop yield recommendation system using machine learning for digital farming. *International Journal of Modern Agriculture*, 10(1), 906-914.

[15] Elavarasan, Dhivya and P. M. Durairaj Vincent. "Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications." *IEEE Access* 8 (2020): 86886-86901.

[16] Rale, N., Solanki, R., Bein, D., Andro-Vasko, J., Bein, W. (2019). Prediction of Crop Cultivation. *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)* 227-232. Springer.

[17] N Balakrishnan, G Muthukumarasamy. "Crop Production-Ensemble Machine Learning Model for Prediction." *International Journal of Computer Science and Software Engineering* 5, no. 7 (2016): 148-153

[18] Dahikar, Snehal S. and S. V. Rode. "Agricultural Crop Yield Prediction Using Artificial Neural Network Approach Miss ." (2014).

[19] Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors*, 18(8), 2674.

[20] <https://www.kaggle.com/siddharthss/crop-recommendation-dataset>

[21] Dahiya, Menal. "A tool of conversation: Chatbot." *International Journal of Computer Sciences and Engineering* 5.5 (2017): 158-161.

[22] Disha Chiplonker, Arjun A., Pruthvi M., Rakshith V. R., Sunil G. L.. Crop yield prediction and alternate crop prediction using ML, *International Journal of Advance Research, Ideas and Innovations in Technology*, www.IJARIIT.com.

[23] <https://data.world/thatzprem/agriculture-india>

[24] Smith, S. L., & Mosier, J. N. (1986). Guidelines for designing user interface software. Bedford, MA: Mitre Corporation.

[25] Colace, F., De Santo, M., Lombardi, M., Pascale, F., Pietrosanto, A., & Lemma, S. (2018). Chatbot for e-learning: A case of study. *International Journal of Mechanical Engineering and Robotics Research*, 7(5), 528-533.

[26] U. Bharti, D. Bajaj, H. Batra, S. Lalit, S. Lalit and A. Gangwani, "Medbot: Conversational Artificial Intelligence Powered Chatbot for Delivering Tele-Health after COVID-19," 2020 5th International Conference on Communication and Electronics Systems (ICCES), 2020, pp. 870-875, doi: 10.1109/ICCES48766.2020.9137944.

[27] White, Judith McIntosh. Source choice in agricultural news coverage: Impacts of reporter specialization and newspaper location, ownership, and circulation. Diss. Texas A & M University, 2010.

[28] Goria, S. (2012). Building website for mobile phone users of an Indian agriculture university library: A model. *DESIDOC Journal of Library & Information Technology*, 32(4).

[29] S. A. Z. Rahman, K. Chandra Mitra and S. M. Mohidul Islam, "Soil Classification Using Machine Learning Methods and Crop Suggestion Based on Soil Series," 2018 21st International Conference of Computer and Information Technology (ICCIT), 2018, pp. 1-4, doi: 10.1109/ICCITECHN.2018.8631943.

[30] Gruyere G, Mehta-Bhatt P, Sengupta D. Bt cotton and farmer suicides in India: reviewing the evidence.

29 and 30 papers not cited





PRIMARY SOURCES

- | | | |
|----------|---|------------|
| 1 | Urmil Bharti, Deepali Bajaj, Hunar Batra, Shreya Lalit, Shweta Lalit, Aayushi Gangwani. "Medbot: Conversational Artificial Intelligence Powered Chatbot for Delivering Tele-Health after COVID-19", 2020 5th International Conference on Communication and Electronics Systems (ICCES), 2020 Publication | 2 % |
| 2 | www.int-arch-photogramm-remote-sens-spatial-inf-sci.net Internet Source | 1 % |
| 3 | Submitted to Southern New Hampshire University - Continuing Education Student Paper | 1 % |
| 4 | Submitted to Liverpool Hope Student Paper | 1 % |
| 5 | Submitted to City University Student Paper | 1 % |
| 6 | "Cyber Intelligence and Information Retrieval", Springer Science and Business Media LLC, 2022 | 1 % |

- 7 www.ijariit.com <1 %
Internet Source
- 8 Adewale Amosu, Yuefeng Sun. "Identification of thermally mature total organic carbon-rich layers in shale formations using an effective machine-learning approach", Interpretation, 2021 <1 %
Publication
- 9 Submitted to Bournemouth University <1 %
Student Paper
- 10 library.seg.org <1 %
Internet Source
- 11 Sanoj Kumar Patel, Anil Sharma, Gopal Shankar Singh. "Traditional agricultural practices in India: an approach for environmental sustainability and food security", Energy, Ecology and Environment, 2020 <1 %
Publication
- 12 Maria Pia Del Rosso, Alessandro Sebastianelli, Dario Spiller, Pierre Philippe Mathieu, Silvia Liberata Ullo. "On-Board Volcanic Eruption Detection through CNNs and Satellite Multispectral Imagery", Remote Sensing, 2021 <1 %
Publication

| | | |
|----|--|------|
| 13 | Internet Source | <1 % |
| 14 | www.ijeat.org Internet Source | <1 % |
| 15 | www.hindawi.com Internet Source | <1 % |
| 16 | www.scribd.com Internet Source | <1 % |
| 17 | Submitted to University of Hertfordshire Student Paper | <1 % |
| 18 | Mauricio R. Bellon, Elisabetta Gotor, Francesco Caracciolo. "Assessing the Effectiveness of Projects Supporting On-Farm Conservation of Native Crops: Evidence From the High Andes of South America", World Development, 2015 Publication | <1 % |
| 19 | www.ubuntupit.com Internet Source | <1 % |
| 20 | Submitted to Free University of Bolzano Student Paper | <1 % |
| 21 | hsat.space Internet Source | <1 % |
| 22 | modern-journals.com Internet Source | <1 % |

23

phpcoderblog.wordpress.com

Internet Source

<1 %

24

www.mdpi.com

Internet Source

<1 %

25

shareok.org

Internet Source

<1 %

Exclude quotes On

Exclude matches < 10 words

Exclude bibliography On