International Journal of Research Publication and Reviews

FarmEasy App (Agricultural technology App) Using Machine Learning Algorithm and Data Analytics

Prof. Junaid Mandviwala, Mohammed Arzan Momin, Rohit Nair, Sufyan Patel,

B.E Computer Engineering, Rizvi College of Engineering, Mumbai, India.
mominarzan@eng.rizvi.edu.in, sufyanpatel@eng.rizvi.edu.in, nairrohit@eng.rizvi.edu.in, junaid@eng.rizvi.edu.in

ABSTRACT

Agriculture is the backbone of Indian economy. Due to global warming and climate change traditional farming in the regular months have been distorted and crops have been ruined is the most common phrase seen today. This not only gives economic losses but also the main reason for farmer suicide. Now agriculture needs support, time has come for technology to take over change. For a crop to grow, favorable soil conditions, ambient rainfall and temperature is necessary. So as now due to climate change temperature and rainfall cannot be well defined, example rains in December and January or irregular temperatures have made it difficult for farmers and common man to predict months of plantation and yield of the crop due to irregularities. So we have formulated an analysis by prediction of a favorable crop based on temperature and current rainfall with soil conditions. Data science in agriculture is a growing field and has a wide scope in future.

Following are various applications of Data Science in agricultural sector:

1. Soil and Crop analysis

2. Weather Prediction

3. Fertilizer Recommendation

4. Disease Detection and Pest Management

5. Adaptation to climate change

6. Automated Irrigation System

But our project focus on Crop analysis, fertilizer detection and diseases detection

Keywords: Crop, Science, Machine Learning,Farming, Data Science, Analysis

Introduction :

within the future, AI will help farmers to transform into agricultural experts, using records to growth yields up to each crop line. AI and ML corporations are growing algorithms that could easily carry out a couple of tasks inside the subject of agriculture. these varieties of algorithms are specialised in disease control and crop harvesting at better speeds than in humans. those sorts of algorithms are trained to evaluate soil high-quality and stumble on weeds by using selecting and packing flowers on the equal time. these algorithms are also able to addressing the demanding situations facing agricultural people. Open AI insect detection machine: insects are one of the worst enemies of farmers to damage plants.

The goal of the venture is to determine and decide the nature and quality of domestically based totally soils, considering the extent of toxicity within the present and predicting its destiny fee using the AI version. the main goals of the mission are: -

- we have a restrained wide variety of sources and their proper use is a major concern, even though the proper use of assets impacts our lives.
- further, there may be a shortage of studies information on this area, that can produce massive quantities of farmer statistics. data, lots of facts, gathered through sensible agricultural sensors, e.g. climatic conditions, soil fine, and plant growth. This statistics can be used to track your commercial enterprise reputation in fashionable in addition to personnel overall performance, useful resource performance, and many

others.

- better manipulate of internal methods and, therefore, reduced manufacturing dangers. The ability to foresee the outcome of your manufacturing allows you to plan for a better product distribution.
- Being able to identify any increase troubles within the crop may be able to reduce the hazard of losing your crop.
- increasing commercial enterprise performance through flexible approaches. by using the use of our app, you may routinely carry out more than one techniques during your manufacturing cycle, e.g. which crop is appropriate for planting, fertilizing, or disorder manipulate

progressed product quality and volumes. Get better manipulate of the manufacturing procedure and hold high levels of crop fine and automatic increase capability.

2.Monographic Review:

- **Survey Existing system**

in the present day synopsis of the economic revolution, in which we've got a confined quantity of sources and their right usage is a subject of extremely good difficulty, whether or not it's the utilization of water or utilization of minerals from ores all this in a roundabout way impacts our lives. With the confined availability of assets and accelerated consumption there charges had been rising up and so there sustainable utilization is vital. further, inside the case of Farming where we need to feed a massive variety of clients, any kind of loss at any level proves to be a huge loss to the financial system and the person as nicely. moreover, there's a loss of research statistics on this discipline. the primary purpose is to convey AI and gadget implemented Farming to India, to ample up the technical application of AI and device studying among Farmers, Researchers, and authorities.

Soil Fertilizer detection and crop to grow gadget: The kind of soil and vitamins of soil performs an crucial thing within the form of crop is grown and the great of the crop. due to increasing, deforestation soil nice is degrading and it's tough to decide the pleasant of the soil. A German-primarily based tech begin-up PEAT has developed an AI-primarily based utility known as Planter which can perceive the nutrient deficiencies in soil including plant pests and illnesses through which farmers can also get an idea to apply fertilizer which helps to enhance harvest nice. This app makes use of photo popularity-based totally era. The farmer can capture pix of flora using smartphones. We also can see soil recovery techniques with suggestions and other answers via short films on this application. similarly, hint Genomics is any other gadget mastering-based organization that enables farmers to do a soil analysis to farmers. Such sort of app allows farmers to reveal soil and crop's fitness situations and produce healthy plants with a higher degree of productiveness. Precision Farming and Predictive Analytics: AI packages in agriculture have evolved packages and tools which help farmer's faulty and managed farming through providing them proper steering to farmers about water management, crop rotation, well timed harvesting, and form of crop to be grown, most effective planting, pest attacks, and nutrition control. at the same time as the use of the system mastering algorithms in connection with snap shots captured by using satellites and drones, AI-enabled technologies are expecting climate situations, analyze crop sustainability and examine farms for the presence of illnesses or pests and poor plant nutrition on farms with statistics like temperature, precipitation, wind pace, and sun radiation.

- **System Limitation existing or research gap**

although synthetic intelligence and machine mastering improves the agriculture enterprise in lots of superb approaches, there are numerous concerns concerning the drawing close of ML on employment and the personnel of the agricultural sectors, Agriculture is a $three trillion industry that employs over 1.5 billion human beings, that's a whopping 20% of the arena's populace, field tasks which can be monotonous can be easily automatic this may gradually ensure roles out of date, Human labour will get replaced with the aid of clever robots which could appropriately navigate the space, find and pass agricultural products as well as perform simple and complex field operations. The value of era along with drones has made it unavailable out of doors of the authorities and research our bodies, it is luxurious to shop for the drones, the largest task will be investment internally from the authorities efforts and research establishments. The concept of trusting information and algorithms greater than our very own judgment has its execs and cons. glaringly, we gain from those algorithms, and in any other case, we wouldn't be the usage of them within the first area.those algorithms allow us to automate strategies with the aid of making knowledgeable judgments the use of available information. every now and then, however, this indicates changing a person's activity with an algorithm, which comes with ethical ramifications. moreover, who do we blame if something goes incorrect? same aspect can show up here.The maximum normally mentioned case currently is self-using vehicles — how can we pick how the automobile ought to react in the event of a fatal collision? in the destiny will we must pick which ethical framework we want our self-riding car to follow when we are buying the vehicle?

system studying is enormously effective for sensors and can be used to assist calibrate and correct sensors while related to different sensors measuring environmental variables which includes temperature, pressure, and humidity.

3. Scope of FarmEasy:

The inexperienced Revolution in the course of the Nineteen Fifties and 1960s remarkably drove up the global meals manufacturing around the world, saving 1000000000 people from starvation. The revolution caused the adoption of new technologies like high-yielding sorts (HYVs) of cereals, chemical fertilizers and aggro-chemical substances, better irrigation and mechanization of cultivation strategies.

India observed suite and followed the usage of hybrid seeds, device, fertilizers and pesticides. at the same time as these practices solved the food scarcity problem, they created some problems too in phrases of immoderate use of fertilizers and pesticides, depletion of floor-water, soil degradation and many others. these problems were exacerbated through loss of schooling to apply modern-day generation and recognition approximately the best utilization of chemical substances and many others. consistent with the UN meals and Agriculture corporation, the worldwide population will increase

via 2 billion by means of 2050. With restrained arable land to be had and exponentially increasing mouths to feed, we're now in need of a 2d green Revolution.

Predictive and advice Analytic in our undertaking – AI and device studying can help farmers with the aid of recommending the sowing dates for different vegetation primarily based on climate conditions. ML fashions can also suggest tweaks in cropping patterns to boost yields. the use of the ancient manufacturing records, weather forecasts, seed records, and call for and supply records, ML may be used to forecast the quantity of seed that should be grown to fulfil the growing wishes. ML and Deep studying applications are used to discover capacity defects and nutrient deficiencies in the soil. The algorithms analyses the soil samples and correlate unique foliage patterns with certain soil defects, plant pests and illnesses.

figuring out Plant sicknesses – Crop images are analysed using computer imaginative and prescient technology and segmented into areas like historical past, healthy element and diseased part. The diseased component is then captured and sent to faraway labs for further analysis. similarly, the leaf pix pre-processing help early detection of pest's infestation and allow farmers to act quick and minimize losses.

4. Research Methodology

4.1 Data-set Collection

The dataset containing the soil specific attributes which are collected from Polytest Laboratories soil testing lab, Pune, Maharashtra, India. In addition, similar sources of general crop data were also used from Marathwada University. The crops considered in our model include groundnut, pulses, cotton, vegetables, banana, paddy, sorghum, sugarcane, coriander. The number of examples of each crop available in the training dataset is shown. The attributes considered where Depth, Texture, PH, Soil Color, Permeability, Drainage, Water holding and Erosion. The above stated parameters of soil play a major role in the crop's ability to remove water and nutrients from the soil. For crop growth to their possible, the soil must provide acceptable environment for it.

4.2 Crop Prediction using Ensembling technique

Ensemble is a statistics mining version additionally recognised as the model Combiners that integrate the energy of two or more models to gain higher prediction, performance than any of its models ought to reap on my own. In our gadget, we use one of the most acquainted Ensembling method called Majority balloting approach. In the balloting approach any variety of base beginners can be used. There ought to be at the least two base novices. The freshmen are selected in a manner that they're succesful to each other yet being complimentary additionally. better the opposition better is the danger of higher prediction. but it is important for the novices to be complimentary because whilst one or few contributors make an blunders, the opportunity of the ultimate members correcting this error could be high. every learner builds itself into a model. The model gets educated the use of the schooling records set furnished. whilst a new statistics has to be labeled, every version predicts the elegance on its own. sooner or later, the class which is expected by majority of the newcomers is voted to be the magnificence label of the new pattern.

4.3 Algorithms (Learning models)

Machine Learning Algorithm: Different machine learning algorithms are being used in order to make comparisons. The different algorithms used are as follows:

- **Logistic Regression**

Logistic regression is a manner modern modeling the possibility today's a discrete outcome given an enter variable. The maximum common logistic regression fashions a binary final results; some thing which could take two values together with authentic/false, sure/no, and so on. Multinational logistic regression can model situations where there are more than viable discrete outcomes. Logistic regression is a beneficial evaluation approach for class issues, wherein you are trying to determine if a brand new sample fits nice into a class.

- **Decision Tree**

A choice tree is a non-parametric technique present day supervised getting to know modern method. in the course of the process a tree like shape is formed. In this, the datasets is damaged down to construct upon the tree sooner or later. in the end, the ensuing output is a tree like structure with each choice nodes and leaf nodes. selection nodes can either have two or more branches even as the leaf nodes imply the final nodes representing class or regression end result.

- **Naive Bayes**

This classifier has capabilities that are statically impartial to one every other. maximum cutting-edge the different classifiers predict some quantity today's correlation between the functions but Naive Bayes models its one-of-a-kind functions as unbiased functions given its elegance. This implements a restriction on the given information, however in practice naive Bayes have more sophisticated strategies to use and experience a few theoretical help for improving its efficiency.

- **XGBoost**

XGBoost is one of the maximum popular gadget getting to knowmodern algorithm in recent times. no matter the state-of-the-art prediction assignment handy, regression or type. XGBoost is well known to offer better answers than different gadget modern day algorithms. In fact, on the grounds that its inception, it has turn out to be the gadget brand new algorithm to cope with structured information.

velocity and performance: at the start written in C++, it is relatively faster than different ensemble classifiers.

middle algorithm is parallelizable: due to the fact the center XGBoost set of rules is parallelizable it could harness the strength state-of-the-art multi-center computer systems. it is also parallelizable onto GPU's and across networks cutting-edge computer systems making it feasible to teach on very huge datasets as nicely.

continuously outperforms different set of rules strategies: It has shown higher overall performance on a diffusion contemporary machine present day benchmark datasets. extensive kind of tuning parameters: XGBoost internally has parameters for move-validation, regularization, person-described goal features, missing values, tree parameters, scikit-learn like minded API and many others.

- **Support Vector Machine (SVM)**

An SVM that is assist Vector device is an instance modern a supervised machine today's model which has many modern day algorithms that analyzes the records that is used for solving both class and regression troubles. We are given some education samples where every sample is marked such that it belongs to one or different latest the two to begin with given categories. assist Vector model set of rules creates a version in which it allots new samples to any present day the given classes. An SVM represents many examples which are taken as dots in space such that the samples belonging to distinct groups are partitioned with a hole among them..

- **Random Forest**

A random forest is a supervised gadget ultra-modern set of rules this is made out of decision tree algorithms. A random wooded area is a machine studying technique that's used to solve regression and type problems. It state-of-the-art ensemble present day, that is a technique that mixes many classifiers to offer solutions to complicated troubles.
A random wooded area eradicates the constraints state-of-the-art a choice tree set of rules. It reduces the overfitting modern day datasets and will increase precision. It generates predictions with out requiring many configurations in applications (like scikit-research).

4.4 Software Requirements

Python

IDE

Anaconda

Jupyter Notebook

4.5 Hardware Requirements

We find that the following list represents the minimum requirements needed to install End-thought Python 3.9 and associated applications:-

- Modern Operating System:

  - Windows 7 or 10

  - Mac OS X 10.11 or higher, 64-bit

  - Linux: RHEL 6/7, 64-bit (almost all libraries also work in Ubuntu)

- x86 64-bit CPU (Intel / AMD architecture)

- 4 GB RAM

- 5 GB free disk space

4.6 Frameworks Required

1. **Pandas**

  2. **NumPy**

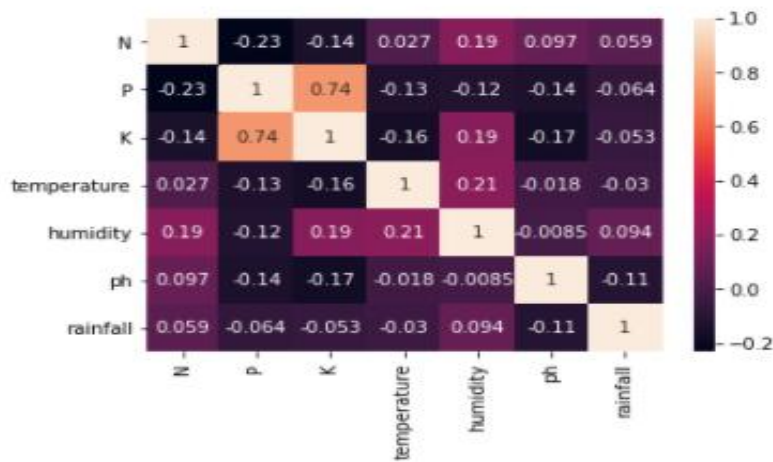 3. **MatPlotLib**

5 Results and Discussion

5.1 Heat Map

```
In [11]:    sns.heatmap(df.corr(),annot=True)

Out[11]:    <matplotlib.axes._subplots.AxesSubplot at 0x7ff83ceb3950>
```


[Figure 1: This photo shows Heat map of this project it is use for probability analysis]

5.2 a) XGBoost Accuracy and Cross Validation score

```
XGBoost's Accuracy is:   0.9931818181818182
                precision      recall    f1-score    support

        apple       1.00        1.00        1.00         13
       banana       1.00        1.00        1.00         17
    blackgram       1.00        1.00        1.00         16
     chickpea       1.00        1.00        1.00         21
      coconut       1.00        1.00        1.00         21
       coffee       0.96        1.00        0.98         22
       cotton       1.00        1.00        1.00         20
       grapes       1.00        1.00        1.00         18
         jute       1.00        0.93        0.96         28
   kidneybeans       1.00        1.00        1.00         14
       lentil       0.96        1.00        0.98         23
        maize       1.00        1.00        1.00         21
        mango       1.00        1.00        1.00         26
     mothbeans       1.00        0.95        0.97         19
      mungbean       1.00        1.00        1.00         24
     muskmelon       1.00        1.00        1.00         23
       orange       1.00        1.00        1.00         29
       papaya       1.00        1.00        1.00         19
    pigeonpeas       1.00        1.00        1.00         18
  pomegranate       1.00        1.00        1.00         17
         rice       0.94        1.00        0.97         16
    watermelon       1.00        1.00        1.00         15

     accuracy                               0.99        440
    macro avg       0.99        0.99        0.99        440
 weighted avg       0.99        0.99        0.99        440
```

[Figure 2: This photo shows XGBoost accuracy over a sample data for analysis of which algorithm shows highest yield in prediction]

5.2 b) Random Forrest Accuracy and Cross Validation Score:

```
RF's Accuracy is:   0.990909090909091
                  precision     recall   f1-score      support

       apple          1.00       1.00       1.00           13
      banana          1.00       1.00       1.00           17
   blackgram          0.94       1.00       0.97           16
    chickpea          1.00       1.00       1.00           21
     coconut          1.00       1.00       1.00           21
      coffee          1.00       1.00       1.00           22
      cotton          1.00       1.00       1.00           20
      grapes          1.00       1.00       1.00           18
        jute          0.90       1.00       0.95           28
  kidneybeans         1.00       1.00       1.00           14
      lentil          1.00       1.00       1.00           23
       maize          1.00       1.00       1.00           21
       mango          1.00       1.00       1.00           26
    mothbeans         1.00       0.95       0.97           19
    mungbean          1.00       1.00       1.00           24
   muskmelon          1.00       1.00       1.00           23
      orange          1.00       1.00       1.00           29
      papaya          1.00       1.00       1.00           19
   pigeonpeas         1.00       1.00       1.00           18
  pomegranate         1.00       1.00       1.00           17
        rice          1.00       0.81       0.90           16
  watermelon          1.00       1.00       1.00           15

    accuracy                                0.99          440
   macro avg          0.99       0.99       0.99          440
weighted avg          0.99       0.99       0.99          440
```

```
# Cross validation score (Random Forest)
score = cross_val_score(RF,features,target,cv=5)
score
```

```
array([0.99772727, 0.99545455, 0.99772727, 0.99318182, 0.98863636])
```

[Figure 3: This photo shows Random Forest accuracy over a sample data for analysis of which algorithm shows highest yield in prediction]

5.3 Results –

[Figure 4: This photo shows the accuracy of which model is the best in dataset]
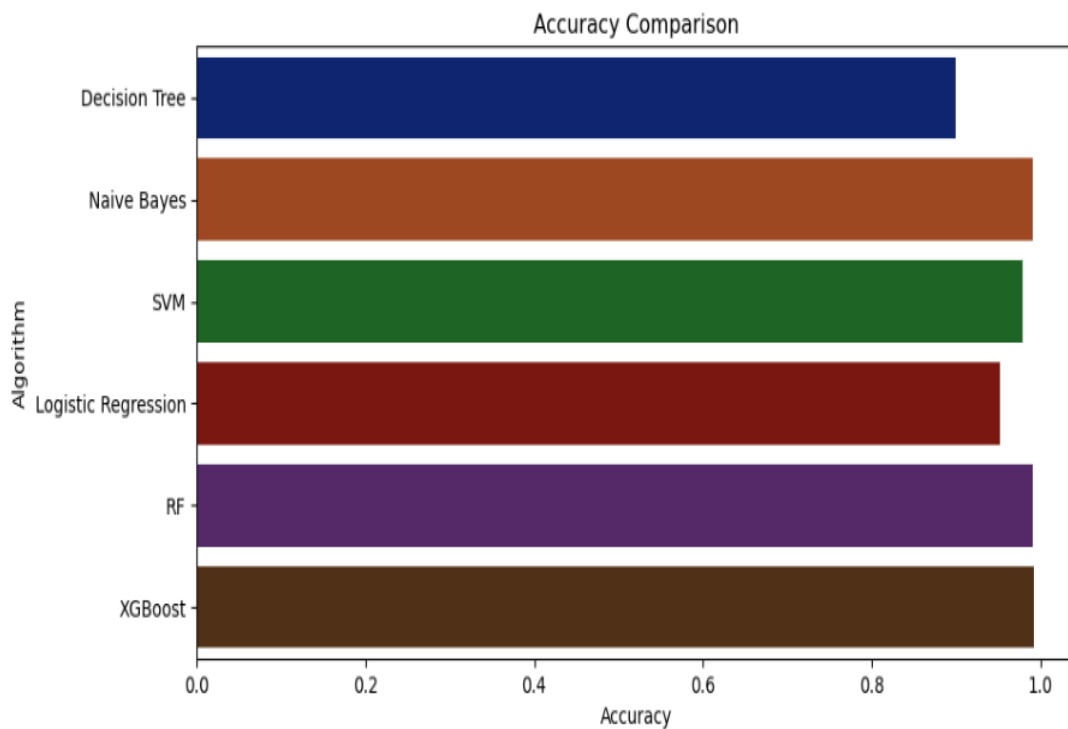
Conclusion:

From above we've discovered that XGBoost presents the satisfactory result and excellent also, but because of its large time taken while in comparison to Random woodland. We are able to use Random Forrest for our undertaking. However this may't be considered as a big stable piece of clue as only some part of the given dataset had been used for this algorithm. The usage of Random woodland we've correctly predicted the Crop appropriate for

given soil parameters, fertilizer detection and additionally in picture processing.so that is fine algorithm for our undertaking

The aim of our challenge is to seize the parameter of the soil exceptional parameters like sodium stage, potassium stage, PH stages, Rainfall, vicinity and so forth. Use this big information for prediction through help of random forest algorithm we will are expecting. The FarmEasy App allows in increasing the accuracy and velocity in the long run obtain the high-precision algorithms for proper prediction and assist farmer in farming. This gadget is designed to decrease the human attempt and ensure that each one hazard are removed. The principle scope of machine is to help farmers.

References:

i. M. Kalimuthu, P. Vaishnavi and M. Kishore, "Crop Prediction using Machine Learning," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020, pp. 926-932, doi: 10.1109/ICSSIT48917.2020.9214190.

ii. P. S. Vijayabaskar, R. Sreemathi and E. Keertanaa, "Crop prediction using predictive analytics," 2017 International Conference on Computation of Power, Energy Information and Commuincation (ICCPEIC), 2017, pp. 370-373, doi: 10.1109/ICCPEIC.2017.8290395.

iii. M. Keerthana, K. J. M. Meghana, S. Pravallika and M. Kavitha, "An Ensemble Algorithm for Crop Yield Prediction," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 963-970, doi: 10.1109/ICICV50876.2021.9388479.

iv. Stuart J. Russell and Peter Norvig, "Artificial Intelligence A Modern Approach "Second Edition" Pearson Education.

v. SarojKaushik "Artificial Intelligence" , Cengage Learning.

vi. George F Luger "Artificial Intelligence" Low Price Edition , Pearson Education., Fourth edition.

vii. P. Louridas and C. Ebert, "Machine Learning," in IEEE Software, vol. 33, no. 5, pp. 110-115, Sept.-Oct. 2016, doi: 10.1109/MS.2016.114.

viii. S. Ray, "A Quick Review of Machine Learning Algorithms," 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), 2019, pp. 35-39, doi: 10.1109/COMITCon.2019.8862451.

ix. Tom M.Mitchell "Machine Learning" McGraw Hill