HINDUSTHAN INSTITUTE OF TECHNOLOGY,COIMBATORE

DEPARTMENT OF INFORMATION TECHNOLOGY

Project Based Experiential Learning Program

(Nalaiyathiran)

FERTILIZER RECOMMENDATION SYSTEM FOR DISEASE PREDICTION

BATCH:2019-2023 **YEAR:**FINAL YEAR

SEM :VII

TEAM MENTOR : Mrs.G.KAVIPRIYA AP/IT

TEAM LEADER : Ms. RAGHA PRIYA S (720819205036)

TEAM MEMBERS : Ms. HARINI K (720819205016)

Ms. PAVITHRA T (720819205031)

Ms.RAGAVI S (720819205035)

TABLE OF CONTENT

ABSTRACT:

Agriculture is the main aspect for the economic development of a country. Agriculture is the heart and life of most species. But in recent days, the field was going down due to various natural calamities. In order to overcome the problem, various issues in this field need to be addressed. The soil type, fertilizer recommendation, diseases in plants and leaves. All these features need to be considered. Our proposed system was organized in such a way, to analyze the soil type, diseases in the leaves and finally to recommend the appropriate fertilizer to the farmers, that may be of great help to them. Plant disease, especially on leaves, is one of the major factors that reduce the yield in both quality and quantity of the food crops. Finding the leaf disease is an important role to preserve agriculture. Smart analysis and Comprehensive prediction models in agriculture helps the farmer to yield the right crop at the right time. The main benefits of the proposed system are as follows: Yield right crop at the right time, Balancing the crop production, control plant disease, Economic growth, and planning to reduce the crop scarcity. Hence to Detect and recognize the plant diseases and to recommend fertilizer it is necessary to provide symptoms in identifying the disease at its earliest. Hence the authors proposed and implemented new fertilizers Recommendation System for crop disease prediction. Detection and recognition of plant diseases using machine learning are very efficient in providing symptoms of identifying diseases at its earliest. Generally, plant diseases are caused by the abnormal physiological functionalities of plants.

CHAPTER 1

INTRODUCTION

1.1 PROJECT OVERVIEW:

Agriculture is the most important sector in today's life. Most plants are affected by a wide variety of bacterial and fungal diseases. Diseases on plants placed a major constraint on the production and a major threat to food security. Hence, early and accurate identification of plant diseases is essential to ensure high quantity and best quality. In recent years, the number of diseases on plants and the degree of harm caused has increased due to the variation in pathogen varieties, changes in cultivation methods, and inadequate plant protection techniques. An automated system is introduced to identify different diseases on plants by checking the symptoms shown on the leaves of the plant. Deep learning techniques are used to identify the diseases and suggest the precautions that can be taken for those diseases.

Agriculture is the main occupation of India. It contributes about sixteen percent (16%) of total GDP and ten percent (10%) of total exports in India. Either directly or indirectly around 60% of people in India depend on agriculture for livelihood. It falls under the primary sector of the Indian economy. It is the main source of food, fodder and fuel. Over 60 % of India's land area is arable making it the second largest country in terms of total arable land. But due to various diseases the quantity and quality of the agricultural product are reducing. Some crop diseases do not have visibility throughout the early stage which results in the damage of the whole crop. Nowadays automatic detection of plant diseases is a very important analysis topic that detects the diseases from the symptoms that seem on the plant leaves. Plant disease is one of the issues that cause reduction in the quality and amount of plant production. This can lead to starvation of peoples. To increase plant productivity and economic process, detection and classification of plant diseases are necessary tasks. It is necessary to detect disease and spray pesticides properly on crops. When they are infected by diseases, there is a change in shape, size and color. These symptoms can be checked manually but not in the proper amount. Hence various image processing methods detect diseases on plant leaves and stems. Using image processing techniques the exact level of disease can be identified based on color, texture or shape change of plants. And then using machine learning algorithms we can classify these diseases and provide a solution for that disease.

1.2.PURPOSE

In India, The Agriculture industry is extremely vital and crucial for economic and social development and jobs. Food demand is increasing due to the rise in population. Most of the Indian population depends on agriculture for their livelihood. To produce food more sustainably farmers are using new technologies.According to “The Economist ", farmers are being “teched up” when it comes to growing crops more sustainable and profitable. It is often heard that diseases and pests attack crops and therefore food gradually reduces due to these attacks. By 2050, the population of earth is expected to grow 9.7 billion. Since, a clear graph of rise in food demand is visible. Agriculture requires a continuous and sustainable increase in productivity, while resources like water, energy, fertilizers etc. need to be used thoughtfully in order to protect and sustain the environment and the soil quality of the arable land.This method leads to the outcome of wrong crop selection is less yield and less profit.

CHAPTER:2

LITERATURE SURVEY

2.1.EXISTING SYSTEM:

The existing method for plant disease detection is simply naked eye observation by experts through which identification and detection of plant diseases is done. For doing so, a large team of experts as well as continuous monitoring of plants is required, which costs very high when we do with large farms. At the same time, in some countries, farmers do not have proper facilities or even the idea that they can contact experts. Due to which consulting experts even cost high as well as time consuming too. In such conditions, the suggested technique proves to be beneficial in monitoring large fields of crops. Automatic detection of the diseases by just seeing the symptoms on the plant leaves makes it easier as well as cheaper. This also supports machine vision to provide image based automatic process control, inspection, and robot guidance.

Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm - Melike Sardogan, Adem Tuncer, Yunus Ozen. Early disease detection is critical in agriculture for efficient crop yield. The diseases bacterial spot, late blight, septoria leaf spot, and yellow curved leaf have an impact on tomato crop quality. Automatic plant disease classification methods also aid in taking action after detecting symptoms of leaf diseases. This paper describes a Convolutional Neural Network (CNN) model and Learning Vector Quantization (LVQ) algorithm-based method for detecting and classifying tomato leaf disease. The dataset contains

500 images of tomato leaves with four disease symptoms. We created a CNN model for automated feature extraction and classification. Color information is being extensively employed in plant leaf disease research .The filters in our model are applied to three channels depending on RGB components.

2.2 REFERENCES:

[1] Reyes Angie .K, Juan C. Caicedo, and Jorge E. Camargo, "Fine-tuning Deep Convolutional Networks for Plant Recognition", In CLEF (Working Notes), 2015.

[2] Hamrouni .L, Aiadi .O, Khaldi .B and Kherfi .M.L, "Plants Species Identification using Computer Vision Techniques", Revue des Bioressources 7, no. 1, 2018.

[3] Dimitrovski, Ivica, GjorgjiMadjarov, DragiKocev, and PetreLameski, "Maestra at LifeCLEF 2014 Plant Task: Plant Identification using Visual Data", In CLEF (Working Notes), pp. 705-714, 2014.

[4] Naresh, Y. G., and H. S. Nagendraswamy, "Classification of medicinal plants: an approach using modified LBP with symbolic representation", Neurocomputing 173, pp: 1789-1797, 2016.

[5] Sue Han, CheeSeng Chan, Paul Wilkin, and Paolo Remagnino, "Deep-plant: Plant identification with convolutional neural networks", In Image Processing ICIP), 2015 IEEE International Conference on, pp. 452-456, IEEE, 2015.

[6] Kaur, Lakhvir, and Vijay Laxmi, "A Review on Plant Leaf Classification and Segmentation", International Journal Of Engineering And Computer Science 5, no. 8, 2016.

[7] Kadir, Abdul, Lukito Edi Nugroho, AdhiSusanto, and Paulus InsapSantosa, "Leaf classification using shape, color, and texture features", arXiv preprint arXiv:1401.4447, 2013.

[8] Lee, Sue Han, CheeSeng Chan, Simon Joseph Mayo, and Paolo Remagnino, "How deep learning extracts and learns leaf features for plant classification", Pattern Recognition 71, pp: 1-13, 2017

[9] Joly, Alexis, Hervé Goëau, HervéGlotin, ConcettoSpampinato, Pierre Bonnet, Willem-Pier Vellinga, Julien Champ, Robert Planqué, Simone Palazzo, and Henning Müller, "LifeCLEF 2016: multimedia life species identification challenges", In International Conference of the Cross-Language Evaluation Forum for European Languages, pp. 286- 310, Springer, Cham, 2016.

[10] Zeiler, Matthew D., and Rob Fergus, "Visualizing and understanding convolutional networks", In European conference on computer vision, pp. 818-833. Springer, Cham, 2014.

[11] Satnam Singh and Manjit Singh Bhamrah, "Leaf identification using feature extraction and neural network", IOSR Journal of Electronics and Communication Engineering 5, pp: 134-140, 2015.

[12] Vijayashree .T and Gopal .A, "Authentication of Leaf Image Using Image Processing Technique", ARPN Journal of Engineering and Applied Sciences 10, no. 9, pp: 4287-4291, 2015.

[13] Nguyen, ThiThanhNhan, ThiLan Le Van Tuan Le, Hai Vu, NataponPantuwong, and Yasushi Yagi, "Flower species identification using deep convolutional neural networks", 2015.

[14] Mohanty, Sharada P., David P. Hughes, and Marcel Salathé, "Using deep learning for image-based plant disease detection", Frontiers in plant science 7, pp: 1419, 2016.

[15] Chaulagain, Basanta, Bhuwan Bhatt, BipinKhatiwada, and BishalChaulagain. "Plant Leaf Recognition", 2013

[16] Pujari, Devashish, Rajesh Yakkundimath, and Abdulmunaf S. Byadgi. "SVM and ANN based classification of plant diseases using feature reduction techniques." IJIMAI 3, no. 7 (2016): 6-14.

[17] Lee, Sue Han, Chee Seng Chan, and Paolo Remagnino. "Multi-organ plant classification based on convolutional and recurrent neural networks." IEEE Transactions on Image Processing 27, no. 9 (2018): 4287-4301.

[18] Metrics for Performance Measurements: https://www.mathworks.com/matlabcentral/answers/418986-how-tocalculate-true-positive-true-negative-false-positiveand-false-negative-as-we-have-segment.

19] Akbarzadeh, Saman, AraPaap, SelamAhderom, BeniaminApopei, and Kamal Alameh. "Plant discrimination by Support Vector Machine classifier based on spectral reflectance." Computers and electronics in agriculture 148 (2018): 250-258.https://www.sciencedirect.com/science/article/pii/ S0168169917310268.

[20] Ramesh, S., and D. Vydeki. "Application of machine learning in the detection of blast disease in South Indian rice crops." Journal of Phytology (2019).

[21] Yalcin, Hulya, and Salary Razavi. "Plant classification using convolutional neural networks."In 2016 Fifth International Conference on Agro-Geoinformatics

2.3.PROBLEM STATEMENT DEFINITION:

Agriculture is the most important sector in today’s life. Most plants are affected by a wide variety of bacterial and fungal diseases. Diseases on plants placed a major constraint on the production and a major threat to food security. Hence, early and accurate identification of plant diseases is essential to ensure high quantity and best quality. In recent years, the number of diseases on plants and the degree of harm caused has increased due to the variation in pathogen varieties, changes in cultivation methods, and inadequate plant protection techniques. An automated system is introduced to identify different diseases on plants by checking the symptoms shown on the leaves of the plant. Deep learning techniques are used to identify the diseases and suggest the precautions that can be taken for those diseases. The solution to the problem is Machine learning, which is one of the applications of Artificial Intelligence, is being used to implement the proposed system. Crop recommendation is going to recommend you the best crop you can grow in your land as per the soil nutrition value and along with as per the climate in that region. And recommending the best fertilizer for every particular crop is also a challenging task. And the other and most important issue is when a plant gets caught by heterogeneous diseases that affect less agricultural production and compromises with quality as well. To overcome all these issues this recommendation has been proposed . Nowadays a lot of research and work is being implemented in the smart and modern agriculture domain. Crop recommendation is characterized by a soil database composed of Nitrogen, Phosphorus, potassium. The ensembles technique is used to build a recommendation model that combines the prediction of multiple machine learning. Models to recommend the right crop based on soil value and the best fertilizer to use.

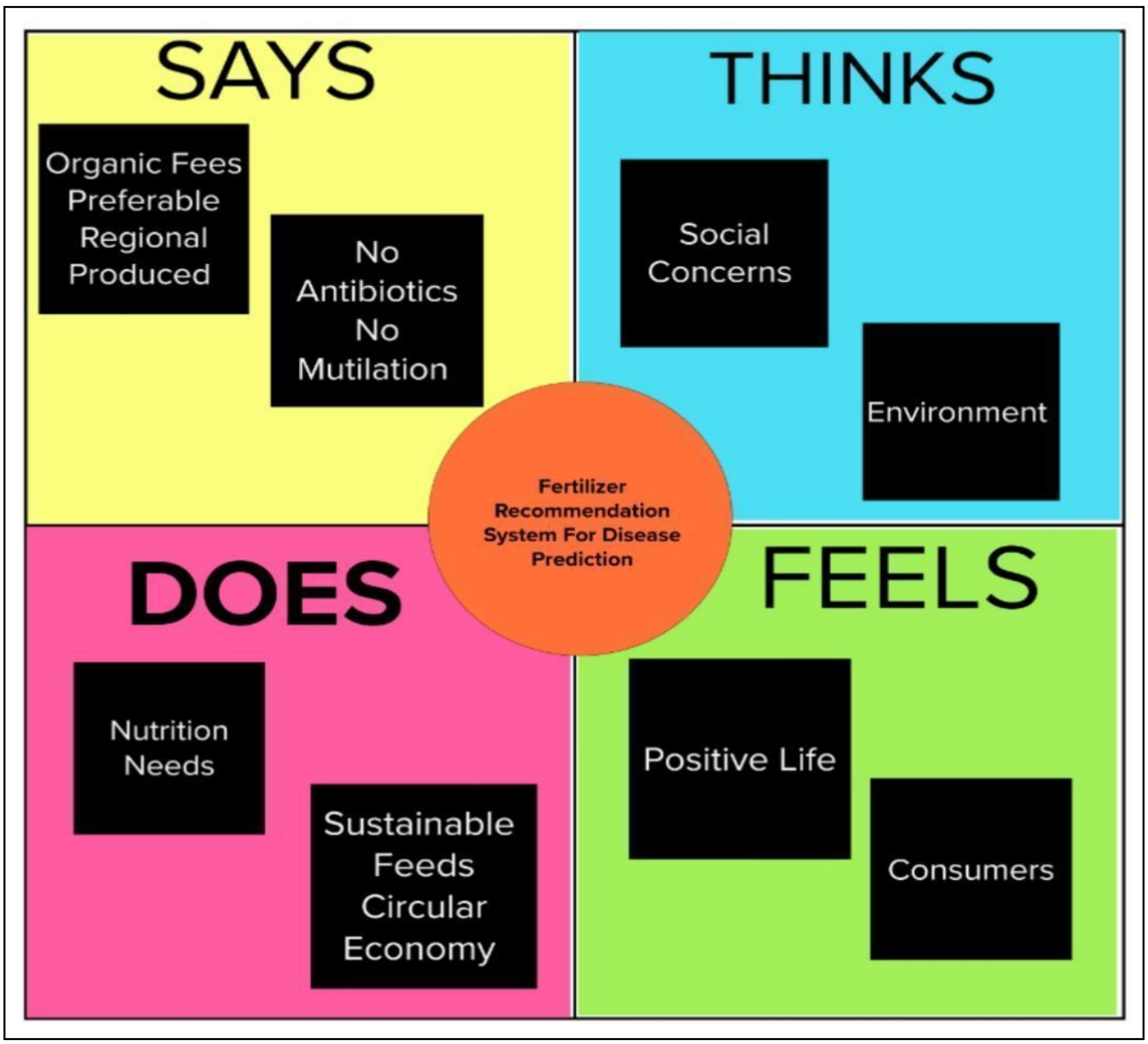
CHAPTER:3

IDEATION & PROPOSED SOLUTION

3.1.EMPATHY MAP CANVAS:

In this diagram, we know about the description of the farmers' thinking.

- ❖ Says.
- ❖ Thinks.
- ❖ Does.
- ❖ Feels.



3.2.IDEATION AND BRAINSTORMING:

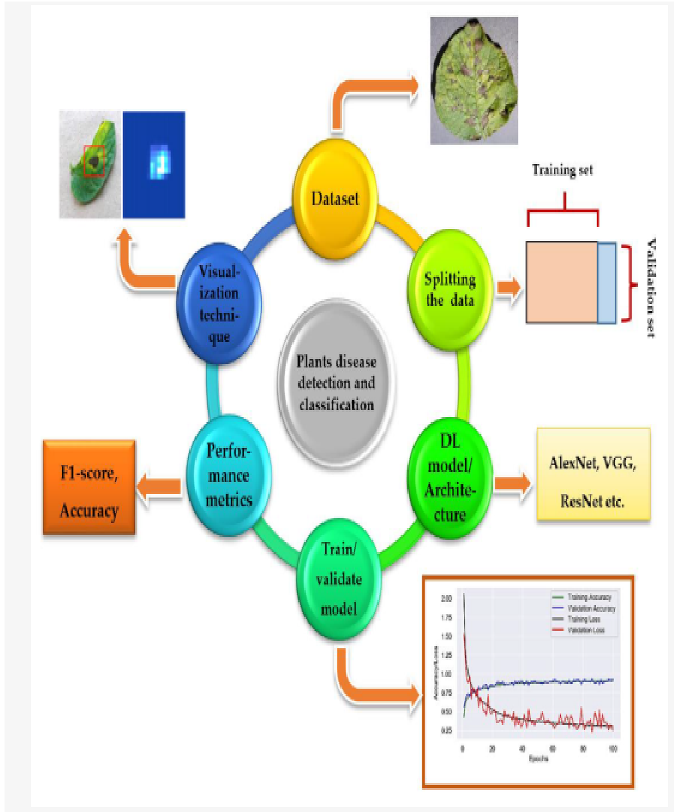
IDEATION:

Detection and recognition of plant diseases using machine learning are very efficient in providing symptoms of identifying diseases at its earliest. Generally, plant diseases are caused by the abnormal physiological functionalities of plants. Therefore, the characteristic symptoms are generated based on the differentiation between normal physiological functionalities and abnormal physiological functionalities of the plants. Mostly, the plant leaf diseases are caused by Pathogens which are positioned on the stems of the plants. These different symptoms and diseases of leaves are predicted by different methods in machine learning. These different methods include different fundamental processes like segmentation, feature extraction and classification and so on. Mostly, the prediction and diagnosis of leaf diseases are dependent on the segmentation such as segmenting the healthy tissues from diseased tissues of leaves.

Leaves are affected by bacteria, fungi, viruses, and other insects. Support Vector Machine (SVM) algorithm classifies the leaf image as normal or affected. Vectors are constructed based on leaf features such as color, shape, textures. Then hyper planes are constructed with conditions to categorize the preprocessed leaves and also implement a multiclass classifier, to predict diseases in the leaf . Recommends the fertilizer for affected leaves based on severity level. It may be organic or inorganic. We can store the fertilizers based on disease categorization with severity levels. The measurements of fertilizers suggested based on disease severity.

BRAINSTORMING:

The dataset is collected then split into two parts, normally into 80% of training and 20% of validation set. After that, DL models are trained from scratch or by using transfer learning techniques, and their training/validation plots are obtained to indicate the significance of the models.



3.3. PROPOSED SOLUTION:

The aim of the proposed system is to help farmers to cultivate crops for better yield. The dataset of crop yield is collected from the last 5 years from different sources. There are 3 steps in proposed work.

Crop Yield Prediction: Crop Yield Prediction can be done using crop yield data, nutrients and location data. These inputs are passed to Random Forest and Support Vector Machine algorithms. These algorithms will predict crop based on present inputs.

Fertilizer Recommendation: Fertilizer Recommendation can be done using fertilizer data, crop and location data. In this part suitable crops and required fertilizer for each crop is recommended. SENSOR for pH the hydrogen-ion in the soil is measured using a PH meter, an electric instrument. A voltage test is used by a pH meter to detect hydrogen ion concentration and consequently pH. It is used to determine the soil's acidity or alkalinity.

A solution with more H+ ions will stay acidic, whereas a solution with OH- ions would remain alkaline. The pH of 1 soil is exceedingly acidic, whereas the pH of 14 soil is excessively alkaline. Pure water, being a neutral solution, has a pH of 7. The pH of various soil samples is determined using a pH meter. It is more precise than using pH strips.

fig:pH probe & Sensor

PROPOSED MODEL FRAME:

The proposed system is used to determine the nutrient quantity of soil through NPK Ratio and predict various diseases crops may be infected with. As we know all the nutrients present in the soil but what amount of nutrients are present in the particular field. Every soil has different micronutrients. But to measure the amount of nutrient available in the soil we are going to design a device which will give a proper reading of the micronutrient and that can be used to predict crops, fertilizers and crop diseases. Following are the main objectives of the proposed system Design and develop a microcontroller-based sensor interfacing for reading soil parameters (NPK value). Converting the sensor value which is analog signal to digital signal for further processing Sending all reading to the system using USB ports available. Developing a website application for displaying the result and generating the report. Based on the result give suggestions to improve the quality of soil.

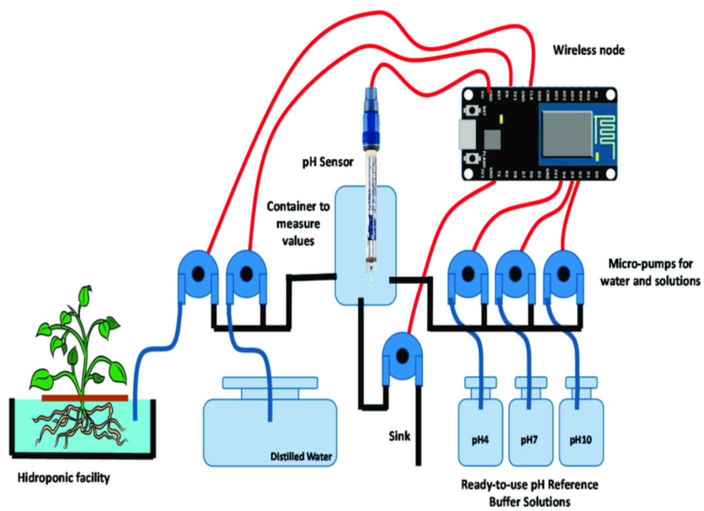
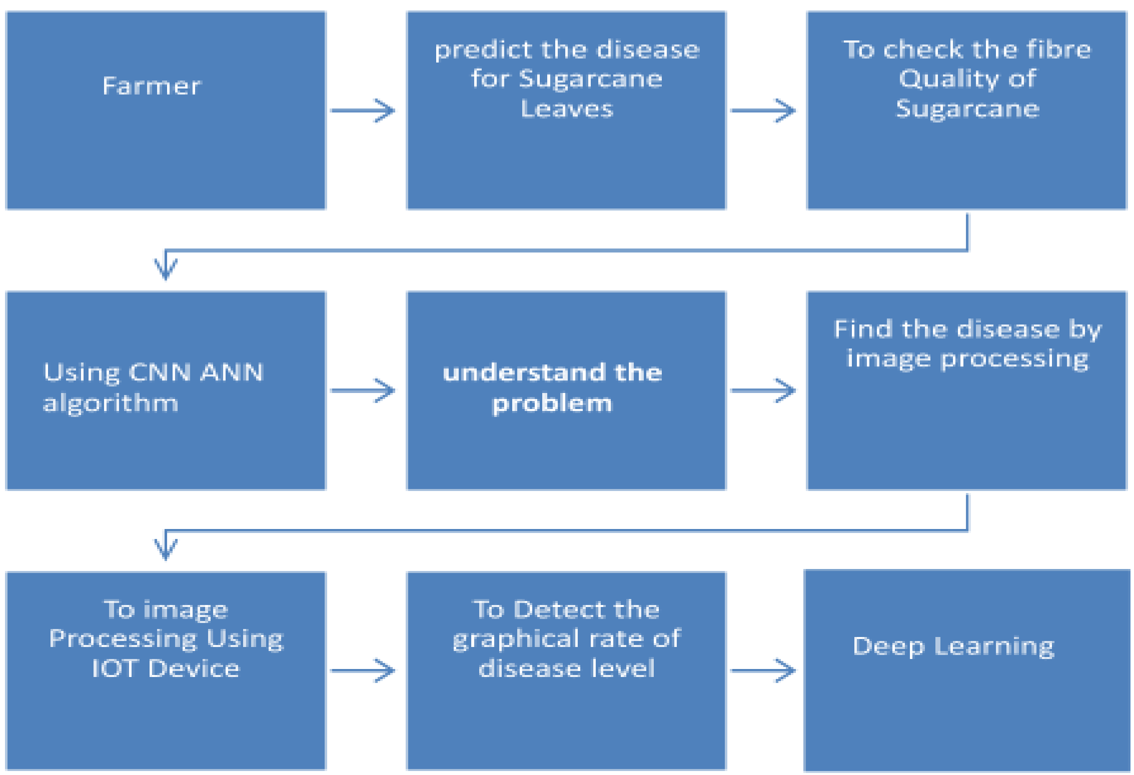


Fig: System design

3.4.PROBLEM SOLUTION FIT:

The solution to the problem is Machine learning, which is one of the applications of Artificial Intelligence, is being used to implement the proposed system. Crop recommendation is going to recommend you the best crop you can grow in your land as per the soil nutrition value and along with as per the climate in that region.

And recommending the best fertilizer for every particular crop is also a challenging task. And the other and most important issue is when a plant gets caught by heterogeneous diseases that affect less agricultural production and compromises with quality as well. To overcome all these issues this recommendation has been proposed . Nowadays a lot of research and work is being implemented in the smart and modern agriculture domain. Crop recommendation is characterized by a soil database composed of Nitrogen, Phosphorus, potassium. The ensembles technique is used to build a recommendation model that combines the prediction of multiple machine learning. Models to recommend the right crop based on soil value and the best fertilizer to use.

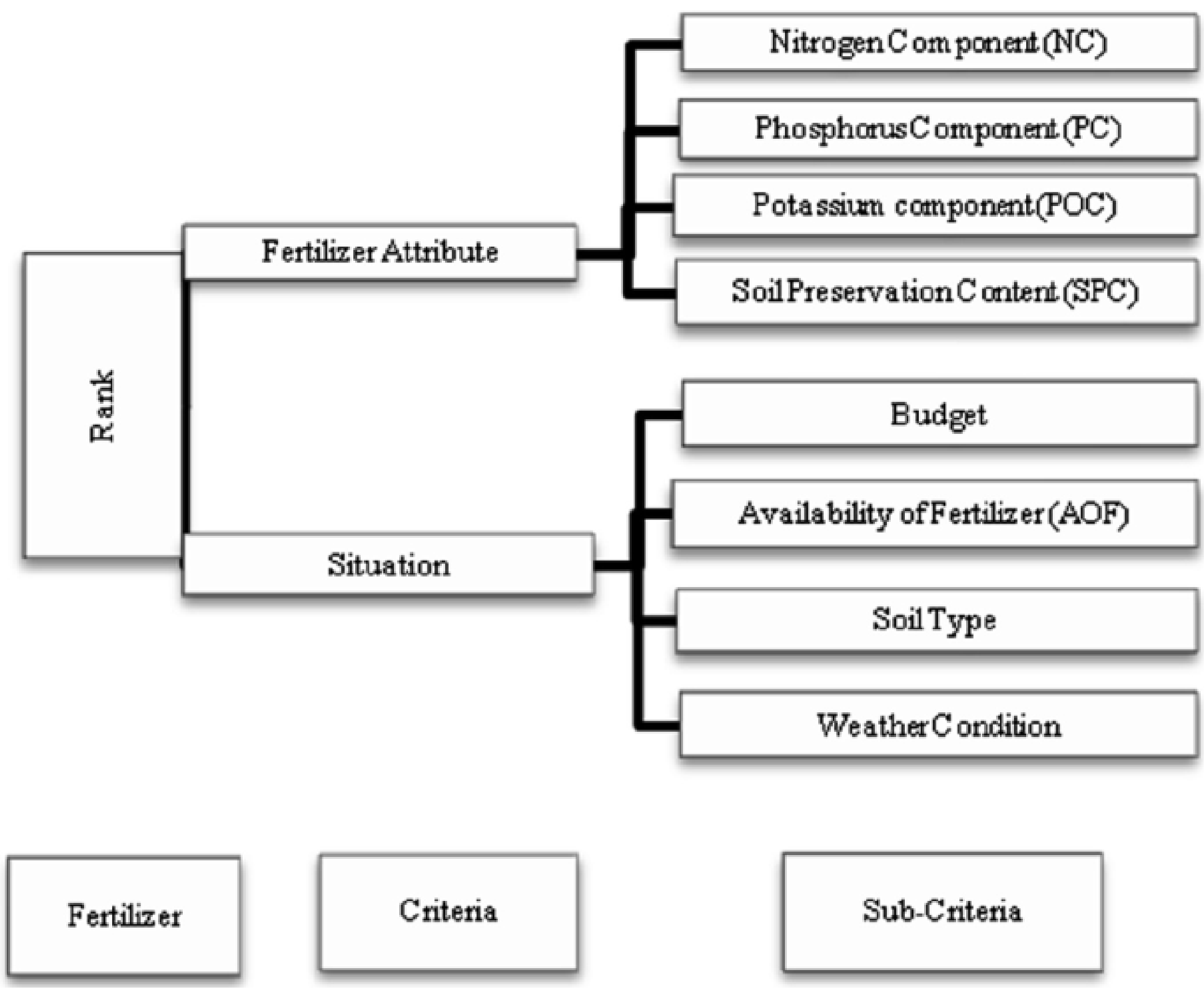


CHAPTER:4

REQUIREMENT ANALYSIS

4.1.FUNCTION REQUIREMENTS:

Statements of services the system should provide, how the system should react to particular inputs and how the system should behave in particular situations. May state what the system should not do.



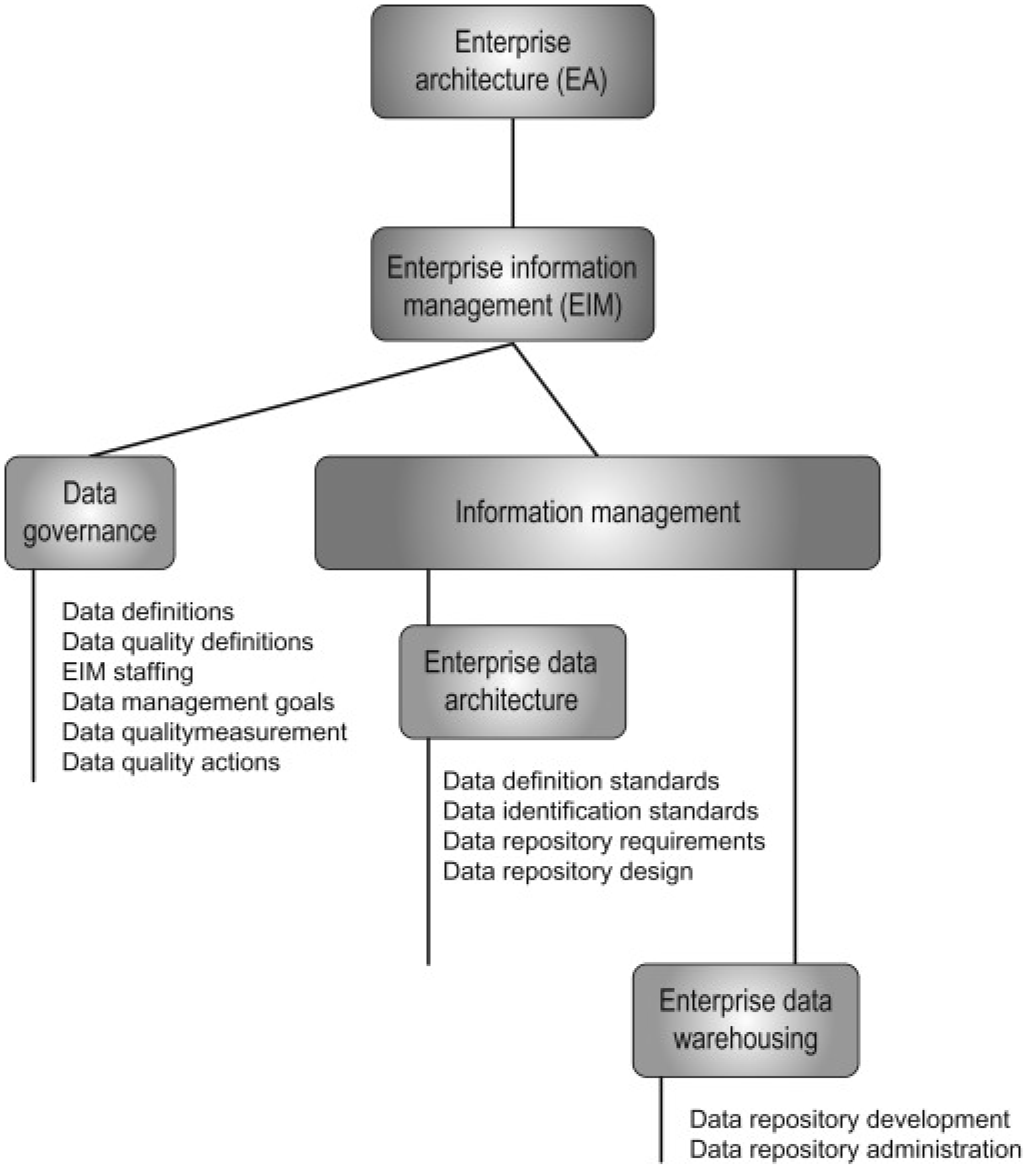
MATERIAL AND METHODS:

A digital camera or similar devices are used to take images of different types, and then those are used to identify the affected area in leaves. Then different types of image-processing techniques are applied to them, the process those images, to get different and useful features needed for the purpose of analyzing later-Plant leaf disease identification is especially needed to predict both the quality and quantity of the First segmentation step primarily based on a mild polygonal leaf model is first achieved and later used to guide the evolution of an energetic contour. Combining global shape descriptors given by the polygonal model with local curvature based features, the leaves are then classified overleaf datasets. In this research work introduce a method designed to deal with the obstacles raised by such complex images, for simple and plant leaves. A first segmentation step based on the graph-cut approach is first performed and later used to guide the evolution of leaf boundaries, and implement classification algorithms to classify the diseases and recommend.

4.2.NON- FUNCTION REQUIREMENT:

Constraints on the services or functions offered by the system such as timing constraints, constraints on the development process, standards, etc.

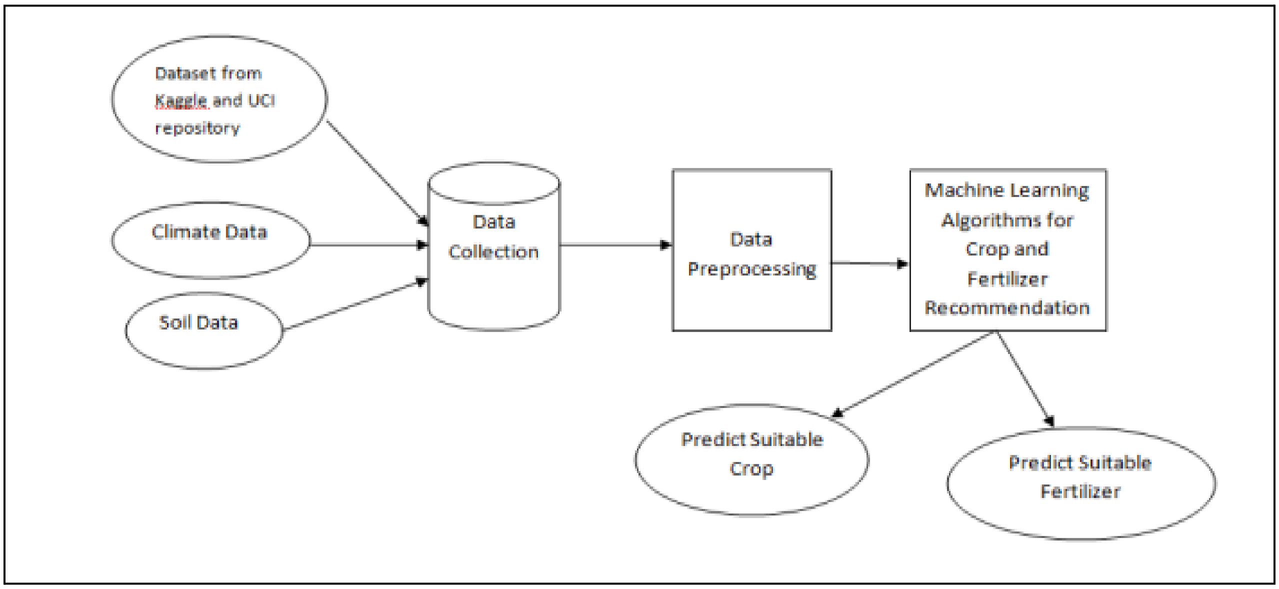
Often apply to the system as a whole rather than individual features or services.


Enterprise architecture (EA)
Enterprise information management (EIM)
Data governance
Information management
Data definitions
Data quality definitions
EIM staffing
Data management goals
Data qualitymeasurement
Data quality actions
Enterprise data architecture
Data definition standards
Data identification standards
Data repository requirements
Data repository design
Enterprise data warehousing
Data repository development
Data repository administration

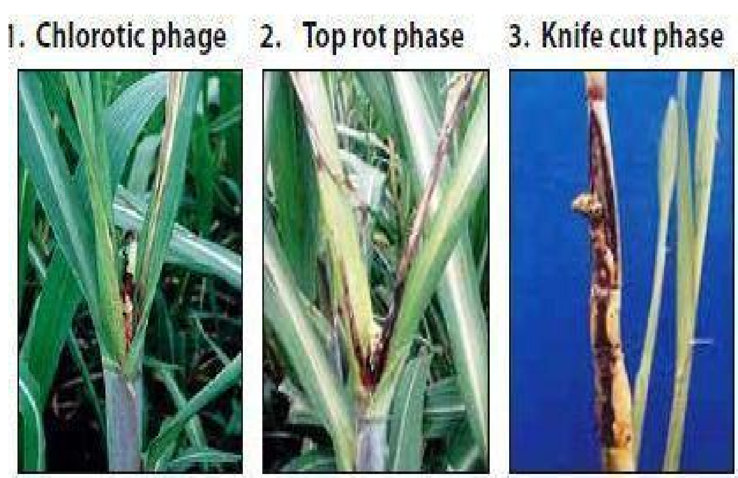
CHAPTER:5

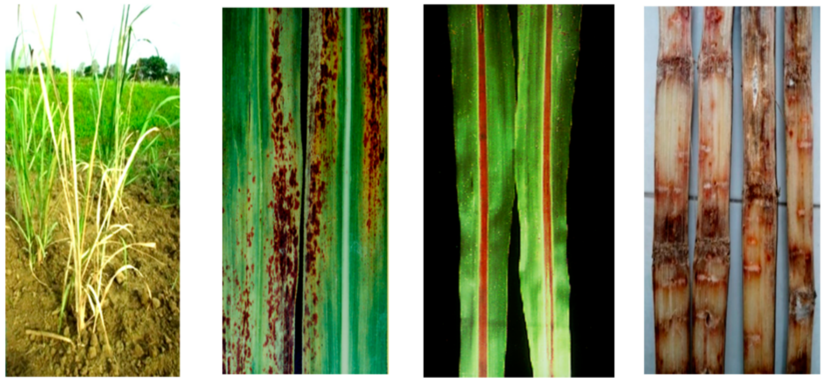
PROJECT DESIGN

5.1.DATA FLOW DIAGRAM:



Data Collection:





(a) Tiller red rot (b) Lamina red rot (c) Midrib red rot (d) Stem red rot

DatasetLink:https://www.kaggle.com/code/emmarex/plant-disease-detection-using-keras

Dataset Splitting:

The historical weather data obtained were divided into two sets, namely the training set and training set. Initially, the artificial intelligence (AI) model is built on a dataset called a training set and the built model is tested on a new set called the test set. The newly developed ML model is applied to test the dataset to measure the performance. This dataset was further split in two different ways. As per the bifurcation of the dataset, if 70–30% split is considered then the data values for the year field that are less than equal to 2010 and greater than equal to 1989 will be considered for training. The data values for the year field that are greater than 2010 and equal to 2019 will be considered for testing. If 80–20% split is considered, then the data values for the year field that are less than equal to 2013 and greater than equal to 1989 will be considered for training and the data values for the year field that are greater than 2013 and equal to 2019 will be considered for testing.

Table 1, shows these two variations in dataset division.

Table 1

Two cases of data split considered in the manuscript for implementation, summary of activation functions, MAE values for two cases of data split for four activation functions, comparison based on overall accuracy of the model for different activation functions.

% of train–test dataset split	Total weather instances	Training instances	Testing instances
70–30%	1,643	1,150	493
80–20%	1,643	1,313	329

5.2.SOLUTION AND TECHNICAL ARCHITECTURE:

Solution Architecture:

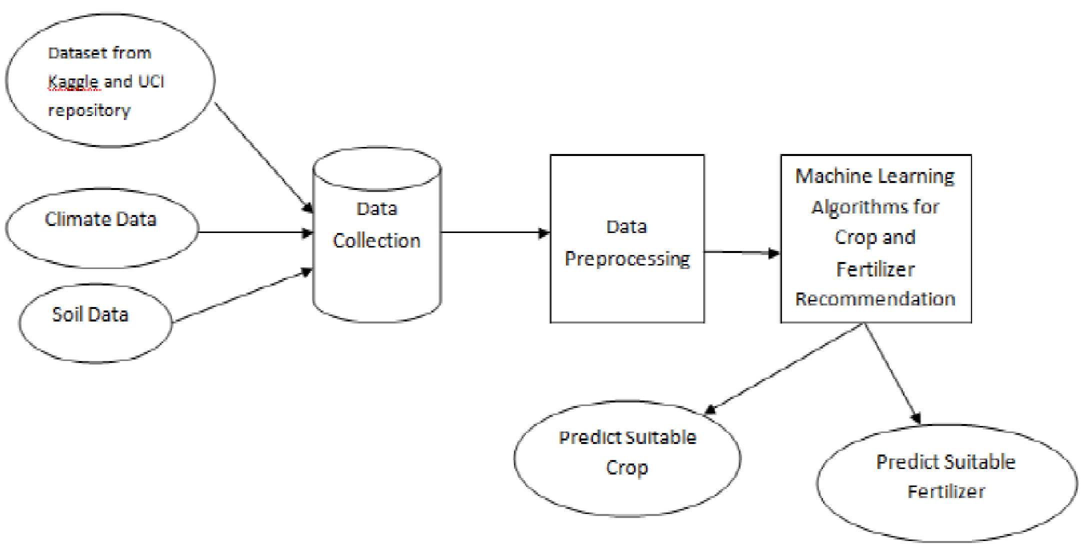


fig: solution architecture

Technical Architecture:

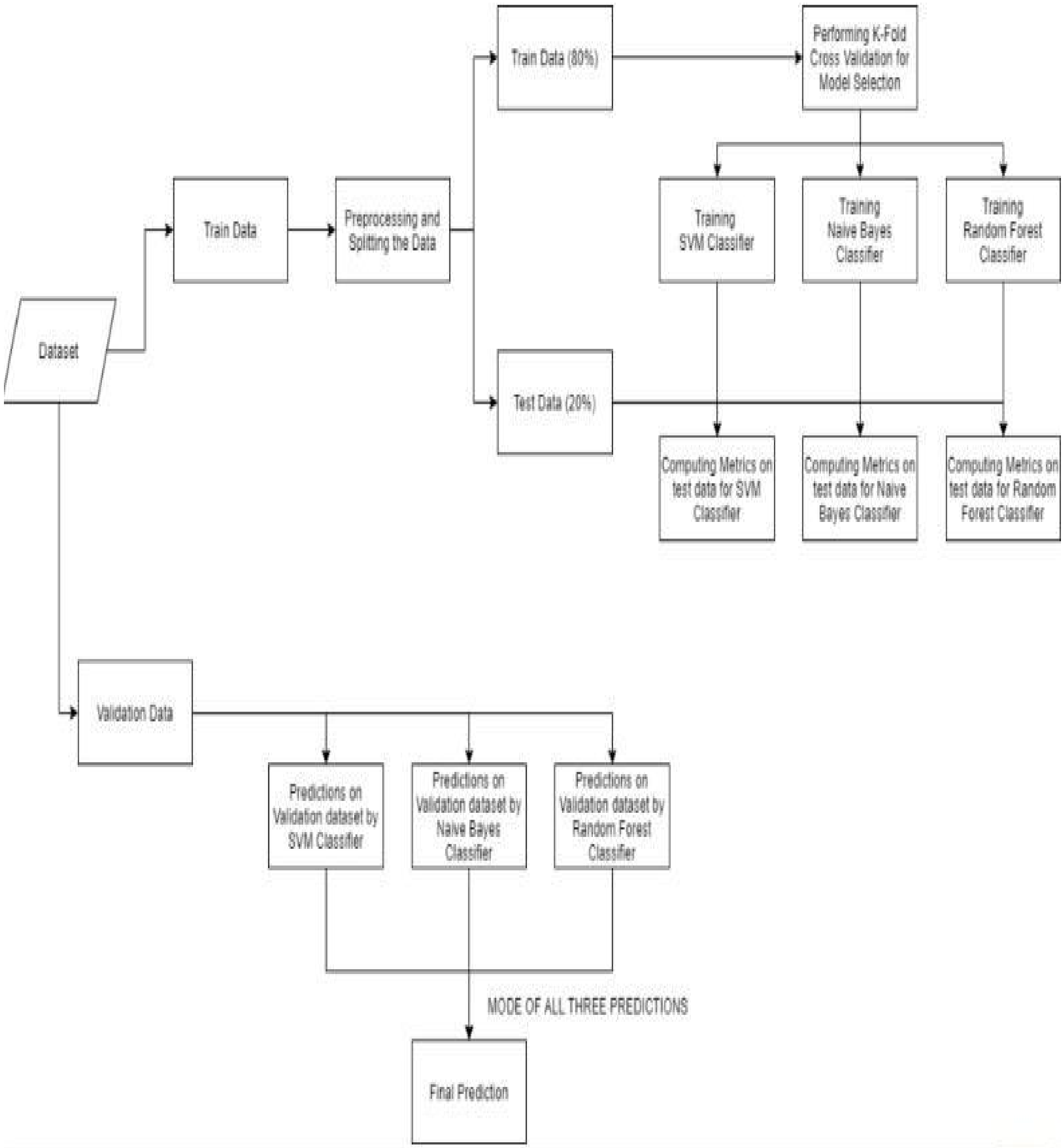


fig: technical architecture

Technology Architecture:

Components & Technologies

S.No	Component	Description	Technology
1	User Interface	user interacts with the application .To depict the human computer interaction and communication.	HTML, CSS,JSP
2	Application Logic 1	A page to upload images as input	Python

5.3.USER STORIES:

- Farmer
- Common People
- Seller
- Buyer
- Employees
- Industrial People

Value for society :

Consumer Farming is one of the major sectors that influences a country's economic growth. In countries like India, the majority of the population is dependent on agriculture for their livelihood. Many new technologies, such as Machine Learning and Deep Learning, are being implemented into agriculture so that it is easier for farmers to grow and maximize their yield

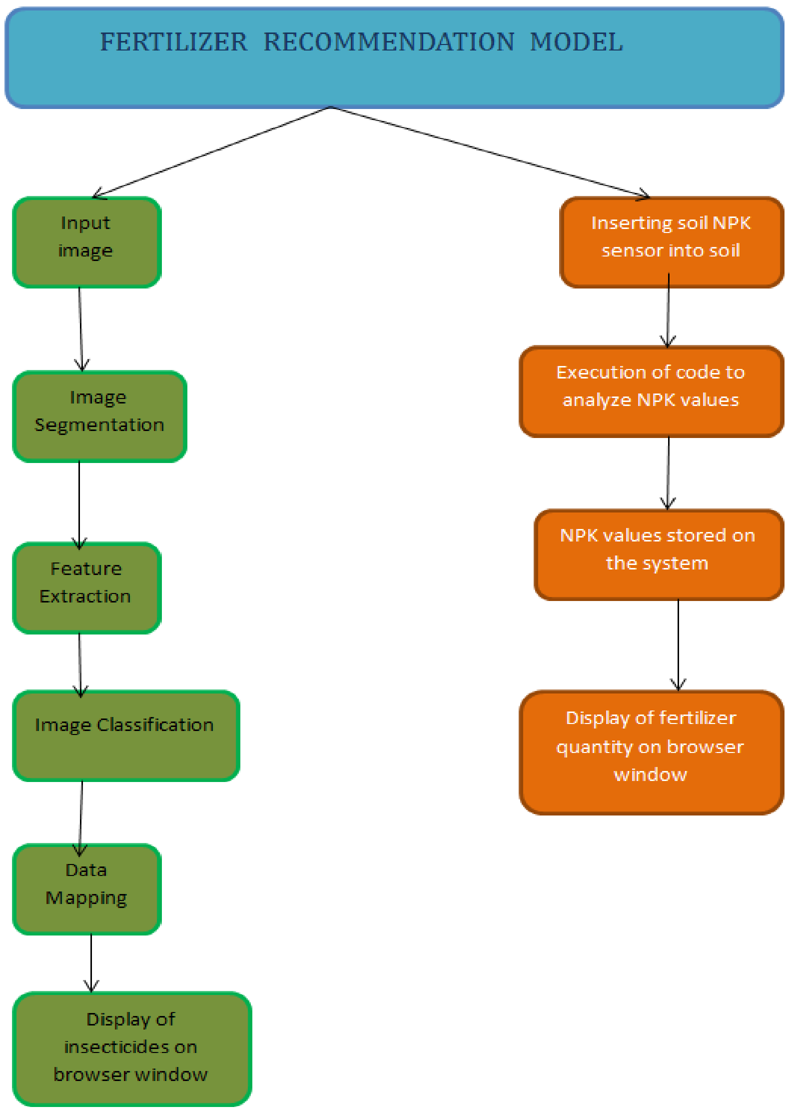
Value for Environment :

- In the crop recommendation application, the user can provide the soil data from their side and the application will predict which crop should the user grow.
- For the fertilizer recommendation application, the user can input the soil data and the type of crop they are growing, and the application will predict what the soil lacks or has excess of and will recommend improvements.
- For the last application, that is the plant disease prediction application, the user can input an image of a diseased plant leaf, and the application will predict what disease it is and will also give a little background about the disease and suggestions to cure it.
- These all are to improve Agriculture, which slightly reduces poverty, climatic conditions, soil erosion etc..

CHAPTER:6

PROJECT PLANNING & SCHEDULING

6.1. SPRINT PLANNING & ESTIMATION:

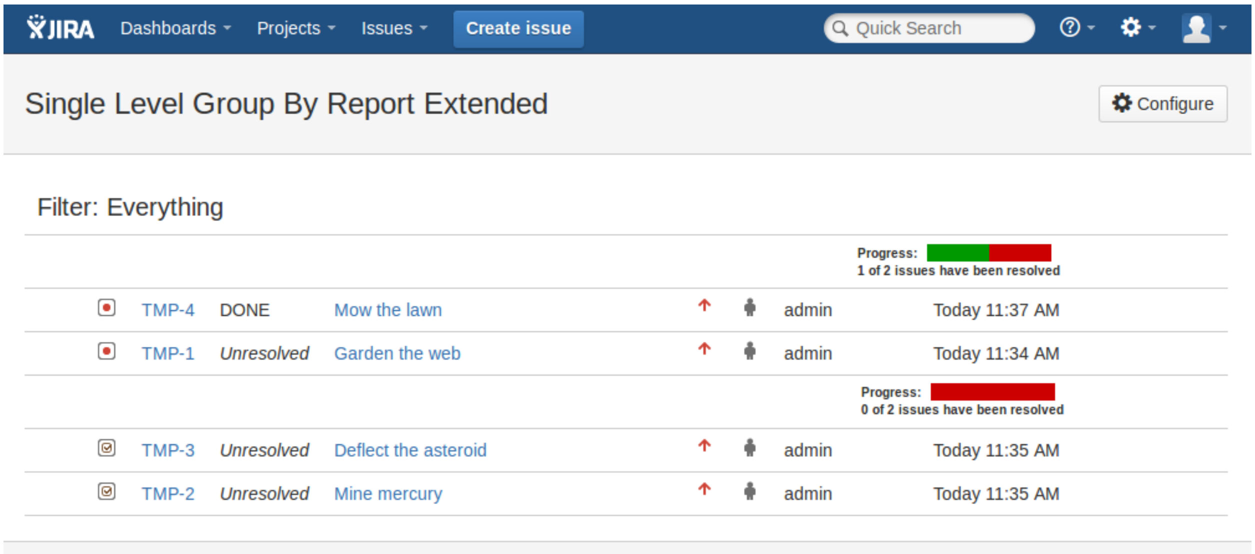


At present, due to the unavailability of natural resources, society should take the maximum advantage of data, information, and knowledge to achieve sustainability goals. In today's world condition, the existence of humans is not possible without the essential proliferation of plants. In the photosynthesis procedure, plants use solar energy.

6.2.SPRINT DELIVERY SCHEDULE:

Crops	Actual	RF	SVM	RF_Err	SVM_Err	RF_Ac	SVM_Ac
Rice	119	120	119	0.41841	0	99.5815	100
Jowar	125	121	125	1.62602	0	98.3739	100
Wheat	126	124	126	0.8	0	99.2	100
Soybeans	110	111	110	0.45249	0	99.5475	100
Sunflower	125	114	125	4.60251	0	95.3974	100
Cotton	146	146	146	0	0	100	100
Sugarcane	98	102	98	2.0	0	98	100
Tobacco	99	119	110	9.17431	5.26315	90.825	94.7368
Onion	147	130	147	6.13718	0	93.862	100
Dry chilli	186	186	186	0	0	100	100
Avg Accuracy						**97.48**	**99.47**

6.3.Report for JIRA:



CHAPTER:7

CODING & SOLUTIONING

7.1. Feature 1:

In this system the micro- controller device is connected to the system through USB ports available and transmits the data from the device to the system. The result is generated from the received data and suggestions are given.

i.CROP RECOMMENDATION:

The NPK value of the soil is calculated using the pH value supplied from the instrument. An API is used to collect temperature and humidity. The pH value, NPK value, temperature, humidity, and rainfall are the characteristics that are used to forecast the best crop to grow in a given place. The crop is predicted using a machine learning model called XGBoost, which has a 99 percent accuracy rate.

ii.FERTILIZER RECOMMENDATION:

Based on the NPK value acquired from the device for a certain soil, a suitable fertilizer is advised for the crop. Proper recommendations for increasing soil fertility are presented (NPK).

iii.DISEASE DETECTION:

The visual data collected from the user is used to detect crop-based illnesses. Deep learning techniques and CNN models are used to forecast if the crop is affected with which disease, and a viable remedy is then offered to the user.

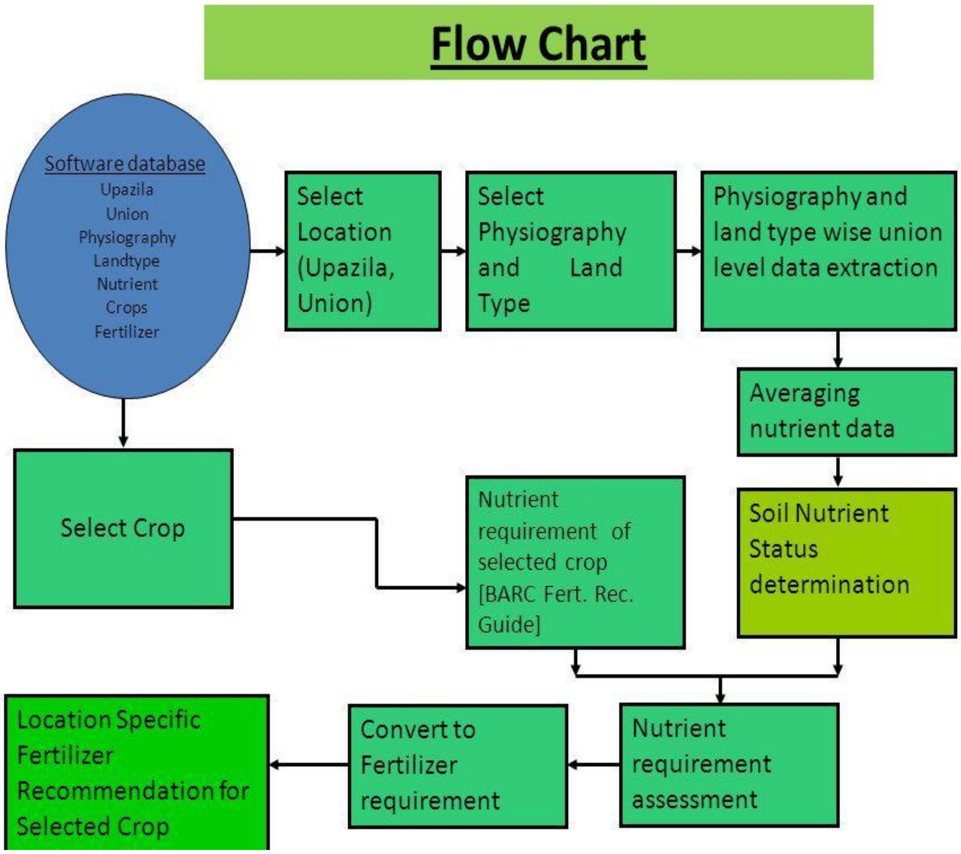


Fig: Flowchart of System

7.2.Feature 2:

The proposed model is a combination of regression and classification. ANN model is used for implementation. Following are the algorithmic steps of the model.

Step 1: Collect week-wise data of agro-meteorological parameters for disease forecasting.

Step 2: Load data by using the pandas library

Step 3: Perform data pre-processing on the instances of the dataset.

Step 4: Build a model based on ANN using Keras.

Step 5: Divide the dataset into two parts: training dataset and test dataset. The dataset is split into two cases (70–30%, 80–30%) and results are verified accordingly.

Step 6: Training the network by utilizing the dataset.

Step 7: Prediction of future values of climatic parameters.

Step 8: Evaluation of the prediction model

Step 9: Prediction of crop disease occurrence. The final output will be in the form of five classes namely Healthy, Rice Blast, Blight, Brown Spot, and False Smut.

Step 10: Evaluation of the classification model.

Class 1—Healthy, Class 2—Rice Blast, Class 3—Bacterial Blight, Class 4—Brown Spot, Class 5—False Smut.

to determine the nutrient quantity of soil through NPK Ratio and predict various diseases crops may be infected with. As we know all the nutrients present in the soil but what amount of nutrients are present in the particular field. Every soil has different micronutrients. But to measure the amount of nutrient available in the soil we are going to design a device which will give a proper reading of the micronutrient and that can be used to predict crops, fertilizers and crop diseases.

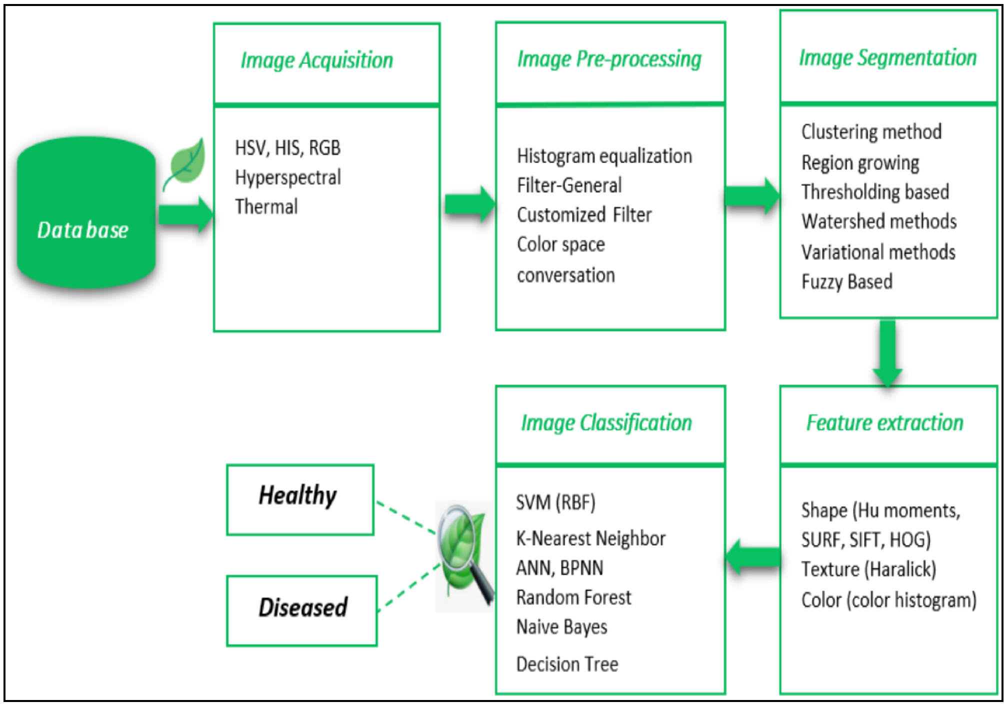
Following are the main objectives of the proposed systemDesign and develop a microcontroller-based sensor interfacing for reading soil parameters (NPK value). Converting the sensor value which is analog signal to digital signal for further processing Sending all reading to the system using USB ports available. Developing a website application for displaying the result and generating the report. Based on the result give suggestions to improve the quality of soil.

CHAPTER:8

TESTING

8.1.TEST CASE:

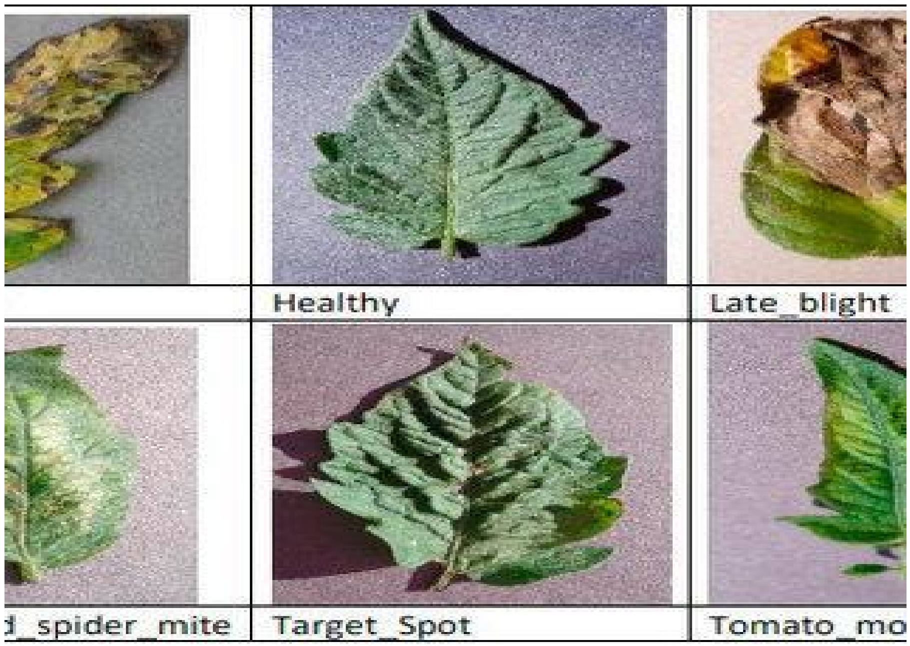
The primary step in identifying diseases is the acquisition of images. In most cases,images can be fetched either from a digital camera or an imaging system. As raw images tend to containoise, removing these impurities is required. As a result, the second step is known as image preprocessing, and involves the removal of unwanted distortions, in addition to contrast enhancement, to clarify and brighten the image features. For example, a Gaussian function that creates soft blur is commonly used to lessen the noise in the image. Subsequently, image segmentation is the third step in which the image is segmented from its background, whereas the region of interest (ROI) is partitioned to emphasize the prominent features. The fourth step is feature extraction [which unveils the information and details of an image. As a side note, the leaf features usually include shape, texture, and color, which are used to diagnose the crop. Thus, these chosen features form an input feature vector which is then fed into the classifier. Using this vector, it is possible to discriminate one class of objects from another. The final step is classification Note that the choice of a suitable classifier depends on the specific problem. The classifier's aim is to recognize the images by sorting them into several predefined classes based on the resulting feature vector obtained in the fourth step. For this purpose, the classification task contains two phases, namely, training and testing. The training operation trains the classifier on a training dataset; thus, the greater the number of training sets, the better the accuracy obtained. It should be noted that the result, which is the crop's healthy state or diseased state associated with the species name, must be achieved as swiftly as possible.



8.2.USER ACCEPTANCE TESTING:

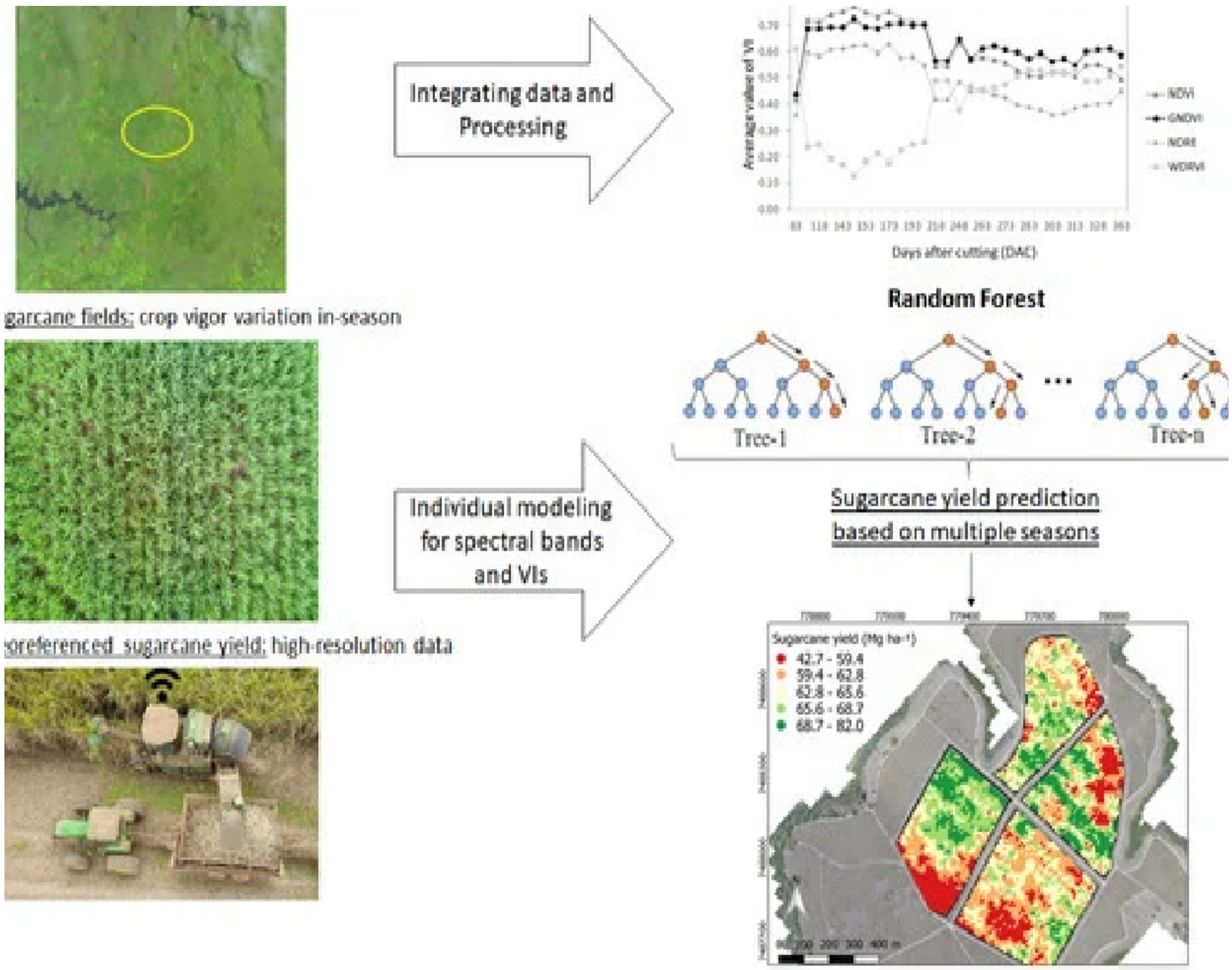
8.2.1.IMAGE CLASSIFICATION:

8.2.1.1.Image acquisition: To get the image of a leaf so that evaluation in the direction of a class can be accomplished.



8.2.1.2Preprocessing:

The purpose of image preprocessing is improving image statistics so that undesired distortions are suppressed and image capabilities which are probably relevant for similar processing are emphasized. The preprocessing receives an image as input and generates an output image as a grayscale, an invert and a smoothed one.

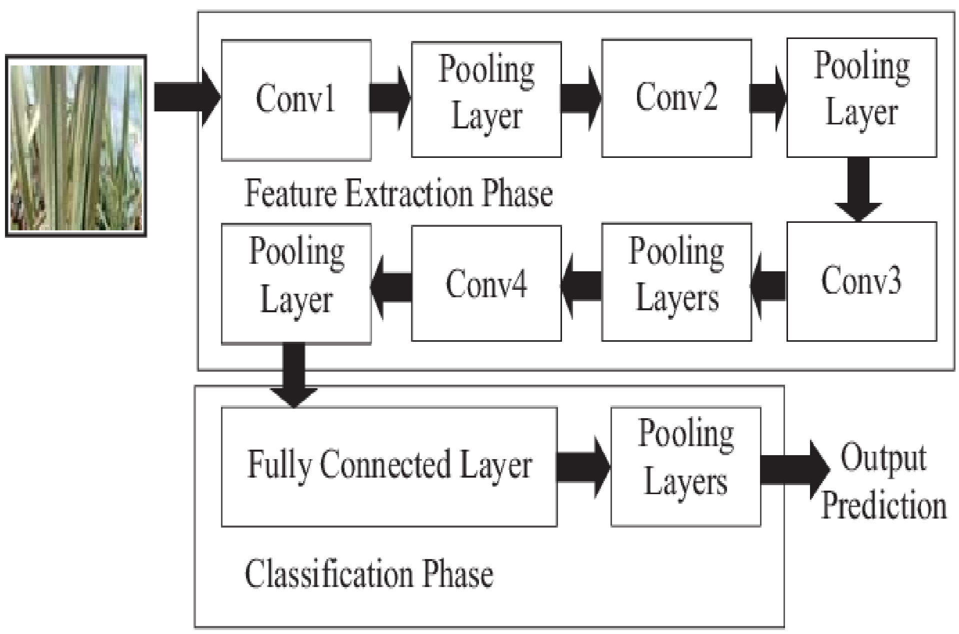


8.2.1.3.Segmentation:

Implements Guided active contour method. Unconstrained active contours applied to the difficult natural images. Dealing with unsatisfying contours, which would try and make their way through every possible grab cut in the border of the leaf. The proposed solution is used the polygonal model obtained after the first step not only as an initial leaf contour but also as a shape prior that will guide its evolution towards the real leaf boundary.

8.2.1.4.Disease Prediction:

Leaves are affected by bacteria, fungi, viruses, and other insects. Support Vector Machine (SVM) algorithm classifies the leaf image as normal or affected. Vectors are constructed based on leaf features such as color, shape, textures. Then hyperplanes are constructed with conditions to categorize the preprocessed leaves and also implement multiclass classifier, to predict diseases in leaf image with improved accuracy.



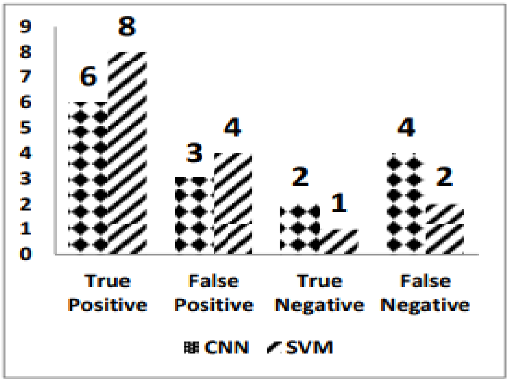
Using CNN algorithm

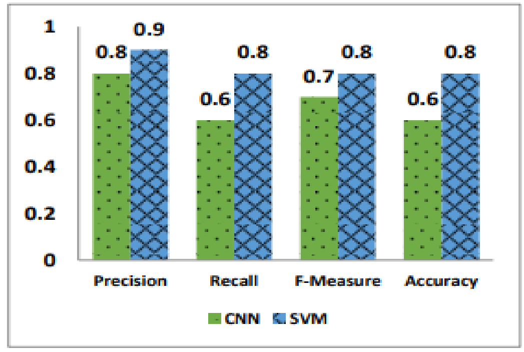
8.2.1.5.Fertilizer Recommendation:

Recommend the fertilizer for affected leaves based on severity level. Fertilizers may be organic or inorganic. Admin can store the fertilizers based on disease categorization with severity levels. The measurements of fertilizers suggested based on disease severity.

8.2.2.SVM Classification Algorithm:

Support Vector Machine(SVM) SVM is a binary classifier to analyze the data and recognize the pattern for classification. The main goal is to design a hyperplane that classifies all the training vectors in different classes. The objective of SVM is to identify a function Fx which obtain the hyper-plane. Hyperplane separates two classes of data sets. The linear classifier is defined as the optimal separating hyperplane. The data sets can be separated in two ways: linearly separated or nonlinearly separated. The vectors are said to be optimally separated if they are separated without error and the distance between the two closest vector points is maximum. For linear separable data sets, training vectors of a different class of pairs (am, bm), where m = 1,2,3,4 …,t am ϵ Rn(Reference Vector) bm ϵ { +1, -1} The decision boundary is placed using a maximal margin between the closest points. w is being a vector perpendicular median to the street. am be the unknown of to be positioned especially elegance according to the decision boundary, and hyperplane (w. a) + c =0 with c as constant For classification (w . am) + c0 $\geq$ 1, $\forall$ bm = +ve samples (1) (w . am) + c0 $\leq$ -1, $\forall$ bm = −ve samples (2) where (w.am) has a dot product of w and am. The inequalities if added i.e multiplying equations (1) and (2) with +1, −1 and bm. Suppose bm such that bm = 1 for +ve samples bm = -1 for -ve samples it results, bm [(w.am) + c0] $\geq$ 1 bm [(w.am) + c0] $\geq$ -1 Therefore rearranging the above equations bm (w.am) + c0 – 1 $\geq$ 0 for points into dataset to in the gutter i.e on the decision boundary bm (w.am) + c0 – 1 = 0.





Performance comparison of CNN and SVM in terms of True Positive, False Positive, True Negative and False Negative

Precision, Recall, F-Measure and Accuracy comparison chart for CNN and SVM

Precision,Recall,E_Measure and Accuracy Value of CNN and SVM :

Classifiers	**Pre**	**Re**	**F-M**	**Acc**
ANN	0.8	0.6	0.7	0.6
SVM	0.9	0.8	0.8	0.8

Advantages :

1) SVM calculation has a regularization parameter, which stays away from over-fitting.

2) SVM calculation utilizes the portion trap, so you can construct master learning about the issue.

3.Random Forest :

Random forest is a supervised machine learning algorithm based on ensemble learning. Ensemble learning is a type of learning where you join different types of algorithms or the same algorithm multiple times to form a more powerful prediction model. The random forest algorithm combines multiple algorithms of the same type. Random Forest algorithms can be used for classification and regression problems.

Advantages:

1) The random forest algorithm is not biased, since there are multiple trees and each tree is trained on a subset of data.

2) Random Forest algorithm is stable if a new data point is introduced in the dataset the overall algorithm is not affected.

CHAPTER -9

RESULT

9.1.PERFORMANCE METRICS:

The complete system is designed using Python. Different datasets like crop, crop yield dataset, Location, soil and crop nutrients, fertilizer datasets are gathered from other sources like agricultural books, agricultural websites.

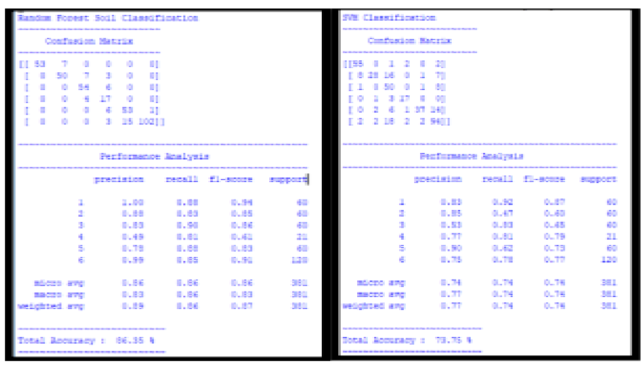


Fig 1. RF and SVM Classification

Figures (Fig 1) shows soil classification using Random Forest algorithm and Support Vector Machine. The output of these algorithms shows confusion matrix as a summary of algorithms different parameters like Precision, Recall averages and accuracy in percentage.

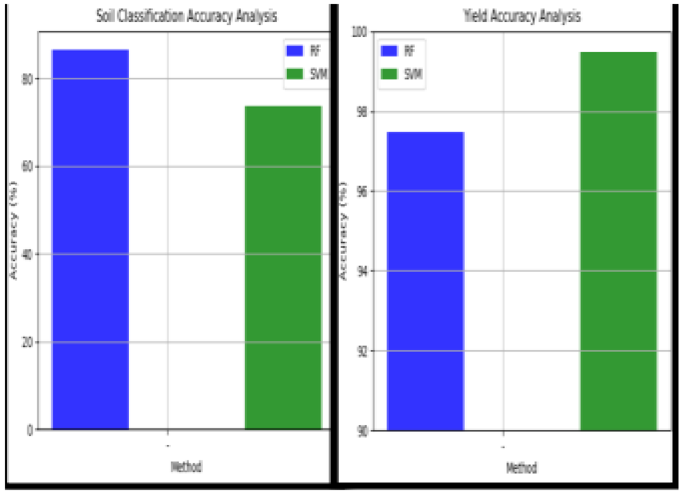


Fig 2.Soil Classification and Crop Yield Analysis

Figures (Fig 2) shows Soil Classification and Crop Yield analysis by graphical representations. The accuracy of Random Forest is 86.35% and Support Vector Machine is 73.75% so Random Forest algorithm is good for Soil Classification. The accuracy of SVM algorithm is 99.47% for yield prediction and RF accuracy is 97.48%. So, for crop yield prediction SVM algorithm is good.

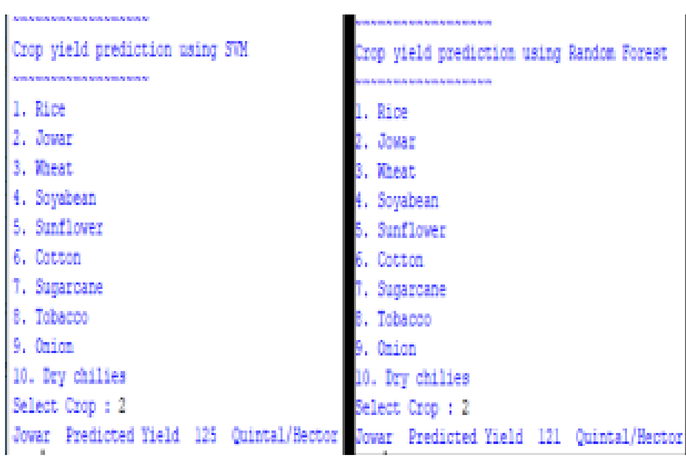


Fig 3.Yield Prediction using RF and SVM

Figures (Fig 3)shows Yield Predicted by both algorithms like if Jowar is selected then Random Forest predicts the “Jowar Predicted Yield 121 Q/H” and SVM predicts “Jowar Predicted Yield is 125Q/H”.

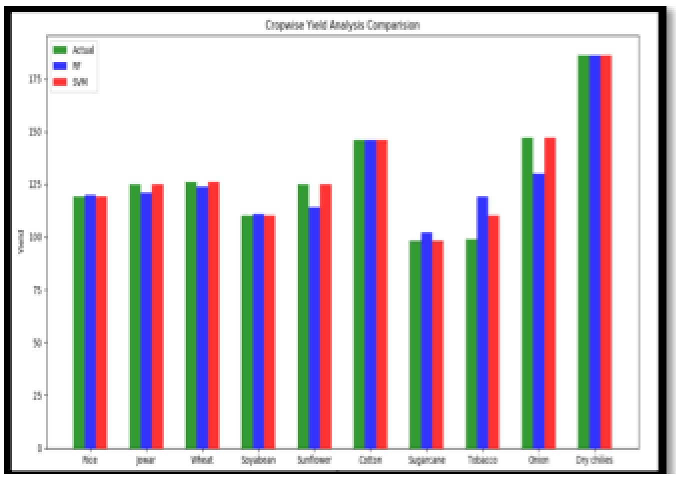


Fig 4.Crop Wise Yield Analysis

Figures (Fig 4)Shows Crop wise yield analysis where each crop yield is taken into consideration.

CHAPTER-10

ADVANTAGES & DISADVANTAGES

To improve the productivity of plants and boost their growth a chemical substance is added to soil or plant tissues is known as Fertilizers. It can be organic or inorganic and both can be used for providing basic nutrients to plants. Most fertilizers consist of nitrogen, potassium, phosphorus and some contain other nutrients like lime, zinc, magnesium, etc. which are necessary for plant growth. It is beneficial for agricultural production. so basically three types of fertilizers are commonly used;

- Chemical
- Organic
- biofertilizers

Advantages of Fertilizers:

Full of nutrients: All three types of fertilizers provide important nutrients to the plants which are helpful in the growth of plants and crops. Potassium, nitrogen, phosphorus are basic elements to boost the yield which is easily provided by all fertilizers in different ways. Chemical fertilizers are made artificially and it has fast production but it boosts quickly the plants so it is used widely whereas organic fertilizers are made by natural process and take time to be prepared but they are good for plants also gives all the necessary nutrient to the plant without harming the nature and human .bio fertilizers are also made by a natural process so it takes time to be processed but it helps soil to fix its problems naturally and convert it into efficient for plants. So in any form fertilizers are the main ingredient of plants that help them to grow properly.

Fast absorption: Sometimes plants need a quick fix to survive, in this type of cases fertilizers play a vital role to improve plants' health. plants need nutrients that can be absorbed quickly which is fulfilled by fertilizers. They are easily soluble and fastly absorbed by plants and as soon as possible it helps to regain and boost plant health.

Easily Available: As the fertilizers are very helpful in fast improvement in plants and best for agricultural production its demand also increases but many factories are working constantly to reach the demand. Thus now all the types of fertilizers are easily available in the market.

Enhance Metabolism: Fertilizers are food for plants that promote their growth. It can only be possible when a proper metabolic activity is processed. Fertilizers are easily digestible by plants and thus increase their metabolism rate to enhance plant growth.

Beneficial for large productions: As the population is increasing, there is a huge demand for food, so good yield is required to fulfill the demand. Here fertilizers become helpful for the good production of crops due to their numerous benefits which promote the fast and healthy growth of plants. For large production, fertilizers become compulsory.

Let us discuss some Disadvantages of Fertilizers:

Expensively: Fertilizers are man-made so they need production in factories which makes them costlier than naturally made manure. But it is important for plant nutrients so it is in demand and thus it has high value.

Over usage can damage plants: Fertilizers are used in moderate quantities if we use excessive fertilizers it surely damages the roots of plants and their tissues and thus plants can die. fertilizers are used according to the need of the plant. Unnecessary use of them can affect the plant's health specially if plants have good fertile soil.

Toxic: There are many types of fertilizers in the market, some of them are chemically made. These chemical fertilizers are harmful to humans and plants also. Skin irritation, respiratory problems commonly occur due to fertilizers. Can pass harmful chemical in our food which affects

Effects of the environment: Indirect or indirect ways fertilizers are affecting the environment. by soil pollution, groundwater pollution and also can affect the growth of plants by faster growth instead of normal. It can affect the ecosystem.

Reduce soil quality: fertilizers can reduce the quality of soil and can harm microorganisms in the soil. Long-term use disturbs the pH of the soil and also reduces the microbial activities which are naturally good for plants.

CHAPTER-11

CONCLUSION

Agriculture assumes an indispensable part in endurance for everybody. Farmers confront several challenges as a result of a variety of uncontrollable factors. As a result, we employ prediction models to overcome the unpredictability of favorable crops or other agriculture-related problems. To forecast the favorable crop, the regression model is employed as a prediction tool. Thus, in this research, linear regression analysis is utilized to create a link between the above-mentioned independent factors and their impacts on crop output, with the goal of increasing agricultural productivity by utilizing the model's favorable crop forecast. The dependence of crop production on temperature is also examined in this paper, and a crop prediction is made based on the results. The crop recommendation system is capable of providing a lasting solution to the problem encountered by farmers. There are some aspects of this study that require additional investigation in order to learn more about crop recommendation prediction using regression analysis. Rainfall, minerals accessible in soil such as potassium, nitrogen, and phosphorus, rainfall, and crop market value may all be included in the input parameters, making the model forecast more accurate and lucrative. Other approaches, such as fuzzy logic and neural networks, might be employed in future studies.

CHAPTER-12

FUTURE SCOPE

This further research is implementing the proposed algorithm with the existing public datasets. Also, various segmentation algorithms can be implemented to improve accuracy. The proposed algorithm can be modified further to identify the disease that affects the various plant organs such as stems and cropsThe vast potential of Indian agriculture remains unexplored, and we still have a long way to go in this field of study, as we need to make the device more compact, lightweight, and inexpensive to farmers. The technology will assist farmers by providing required advice on crops, their growth, and other basic information. It will also offer the location of the nearest store where farmers can purchase fertilizer and other materials. It would also assist farmers in selling their commodities to merchants by providing accurate information on market prices and merchant details. The device can also help farmers calculate crop MSP. The disease detection feature can also be improved by adding dedicated cameras to the device, which will improve the device's accuracy even further.

CHAPTER-13

APPENDIX

SOURCE CODE:

 # Importing libraries

1.from __future__ import print_function

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.metrics import classification_report

from sk learn import metrics

from sk learn import tree

import warnings

warnings.filterwarnings('ignore')

2.df = pd**.**read_csv('../Data-processed/crop-recommendation.csv')

3.df**.**head()

4.df**.**tail()

5.df**.**size

6.df**.**shape

7.df**.**columns

8.df['label']**.**unique()

9.df**.**dtypes

10,sns**.**heatmap(df**.**corr(),annot=**True**)

Separating features and target label

11.features = df[['N', 'P','K','temperature', 'humidity', 'ph', 'rainfall']]
target = df['label']

12.#features = df[['temperature', 'humidity', 'ph', 'rainfall']]
labels = df['label']

13.# Initializing empty lists to append all model's name and corresponding name
acc = []
model = []
# Splitting into train and test data

14.from sklearn.model_selection **import** train_test_split
Xtrain, Xtest, Ytrain, Y test = train_test_split(features,target,test_size = 0.2,random_state =2)

Decision Tree

15.from sklearn.tree **import** DecisionTreeClassifier
DecisionTree = DecisionTreeClassifier(criterion="entropy",random_state=2,max_depth=5)
DecisionTree.fit(Xtrain,Ytrain)
predicted_values = DecisionTree.predict(Xtest)
x = metrics.accuracy_score(Ytest, predicted_values)
acc.append(x)
model.append('Decision Tree')
print("DecisionTrees's Accuracy is: ", x*100)

16.print(classification_report(Ytest,predicted_values))

17.from sklearn.model_selection **import** cross_val_score

18.# *Cross validation score (Decision Tree)*

score = cross_val_score(DecisionTree, features, target,cv=5)

19.score

Saving trained Decision Tree model

20.import pickle

 # Dump the trained Naive Bayes classifier with Pickle

DT_pkl_filename = '../models/DecisionTree.pkl'

 # Open the file to save as pkl file

DT_Model_pkl = open(DT_pkl_filename, 'wb')

pickle**.**dump(DecisionTree, DT_Model_pkl)

 # Close the pickle instances

DT_Model_pkl**.**close()

Guassian Naive Bayes

21.from sklearn.naive_bayes **import** GaussianNB

NaiveBayes = GaussianNB()

NaiveBayes**.**fit(Xtrain,Ytrain)

predicted_values = NaiveBayes**.**predict(Xtest)

x = metrics.accuracy_score(Ytest, predicted_values)
acc.append(x)
model.append('Naive Bayes')
print("Naive Bayes's Accuracy is: ", x)

22.print(classification_report(Ytest,predicted_values))

22.# *Cross validation score (NaiveBayes)*
23.score = cross_val_score(NaiveBayes,features,target,cv=5)
24.score

Saving trained Guassian Naive Bayes model

25.import pickle
# *Dump the trained Naive Bayes classifier with Pickle*
NB_pkl_filename = '../models/NBClassifier.pkl'
# *Open the file to save as pkl file*
NB_Model_pkl = open(NB_pkl_filename, 'wb')
pickle.dump(NaiveBayes, NB_Model_pkl)
# *Close the pickle instances*
NB_Model_pkl.close()

Support Vector Machine (SVM)

26.from sklearn.svm **import** SVC
data normalization with sklearn
from sklearn.preprocessing **import** MinMaxScaler

```
# fit scaler on training data
norm = MinMaxScaler().fit(Xtrain)
X_train_norm = norm.transform(Xtrain)
# transform testing dataabs
X_test_norm = norm.transform(Xtest)
SVM = SVC(kernel='poly', degree=3, C=1)
SVM.fit(X_train_norm,Ytrain)
predicted_values = SVM.predict(X_test_norm)
x = metrics.accuracy_score(Ytest, predicted_values)
acc.append(x)
model.append('SVM')
print("SVM's Accuracy is: ", x)
```

27.

```
print(classification_report(Ytest,predicted_values))
```

28.

```
# Cross validation score (SVM)
score = cross_val_score(SVM,features,target,cv=5)
score
```

29.

```
#Saving trained SVM model
import pickle
# Dump the trained SVM classifier with Pickle
SVM_pkl_filename = '../models/SVMClassifier.pkl'
# Open the file to save as pkl file
SVM_Model_pkl = open(SVM_pkl_filename, 'wb')
pickle.dump(SVM, SVM_Model_pkl)
```

 # Close the pickle instances

SVM_Model_pkl.close()

Logistic Regression

30.from sklearn.linear_model **import** LogisticRegression

LogReg = LogisticRegression(random_state=2)

LogReg.fit(Xtrain,Ytrain)

predicted_values = LogReg.predict(Xtest)

x = metrics.accuracy_score(Ytest, predicted_values)

acc.append(x)

model.append('Logistic Regression')

print("Logistic Regression's Accuracy is: ", x)

print(classification_report(Y test,predicted_values))

31.*# Cross validation score (Logistic Regression)*

score = cross_val_score(LogReg,features,target,cv=5)

32.score

Saving trained Logistic Regression model

33.import pickle

 # Dump the trained Naive Bayes classifier with Pickle

LR_pkl_filename = '../models/LogisticRegression.pkl'

 # Open the file to save as pkl file

LR_Model_pkl = open(DT_pkl_filename, 'wb')

pickle.dump(LogReg, LR_Model_pkl)

 # Close the pickle instances

LR_Model_pkl.close()

Random Forest

34.from sklearn.ensemble **import** RandomForestClassifier

RF = RandomForestClassifier(n_estimators=20, random_state=0)

RF.fit(Xtrain,Ytrain)

predicted_values = RF.predict(Xtest)

x = metrics.accuracy_score(Ytest, predicted_values)

acc.append(x)

model.append('RF')

print("RF's Accuracy is: ", x)

35.print(classification_report(Y test,predicted_values))

36.*# Cross validation score (Random Forest)*

score = cross_val_score(RF,features,target,cv=5)

score

Saving trained Random Forest model

37.import pickle

 # Dump the trained Naive Bayes classifier with Pickle

RF_pkl_filename = '../models/RandomForest.pkl'

 # Open the file to save as pkl file

RF_Model_pkl = open(RF_pkl_filename, 'wb')

pickle.dump(RF, RF_Model_pkl)

 # Close the pickle instances

RF_Model_pkl.close()

XGBoost

38.import x gboost **as** x gb

XB = xgb.XGBClassifier()

XB.fit(Xtrain,Ytrain)

39.predicted_values = XB.predict(Xtest)

40.x = metrics.accuracy_score(Ytest, predicted_values)

acc.append(x)

model.append('XGBoost')

print("XGBoost's Accuracy is: ", x)

41.print(classification_report(Y Test,predicted_values))

42.*# Cross validation score (XGBoost)*

score = cross_val_score(XB,features,target,cv=5)

score

Saving trained XGBoost model

43.import pickle

Dump the trained Naive Bayes classifier with Pickle

XB_pkl_filename = '../models/XGBoost.pkl'

Open the file to save as pk l file

XB_Model_pkl = open(XB_pkl_filename, 'wb')

pickle.dump(XB, XB_Model_pkl)

Close the pickle instances

XB_Model_pkl.close()

Accuracy Comparison

44.plt.figure(figsize=[10,5],dpi = 100)

plt.title('Accuracy Comparison')

plt.xlabel('Accuracy')

plt.ylabel('Algorithm')

sns.barplot(x = acc,y = model,palette='dark')

45.accuracy_models = dict(zip(model, acc))

46.for k, v **in** accuracy_models.items():

    print (k, '-->', v)

Making a prediction

47.data = np.array([[104,18, 30, 23.603016, 60.3, 6.7, 140.91]])

prediction = RF.predict(data)

print(prediction)

48.data = np.array([[83, 45, 60, 28, 70.3, 7.0, 150.9]])

prediction = RF.predict(data): print(prediction)

GitHub & Project Demo Link:

IBM-EPBL/**IBM-Project-30707-1660155718**

https://uploadnow.io/f/t6f3lVZ