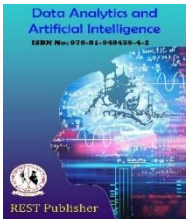
Data Analytics and Artificial Intelligence
Vol: 3(6), 2023
REST Publisher; ISBN: 978-81-948459-4-2
Website: http://restpublisher.com/book-series/daai/



Recent Trends in Detecting Crop Ailments Using Machine Learning Techniques

[*1]S.M.Mustafa Nawaz, [1]K.Maharajan, [2]J. Vijay Anand
[1]*Kalasalingam Academy ofResearch and education, Tamil Nadu, India.*
[2]*J.N.N Institutes of Engineering, Tamil Nadu, India.*
[*]Corresponding author Email: salmannawaz81@gmail.com

Abstract:*This article provides a comprehensive analysis of studies on the diagnosis of crop ailments using machine learning techniques (MLT). Agriculture is the major resource for a densely populated country like India. Farmers faced numerous challenges with the traditional method, including unpredictable weather, plant diseases, and water shortages, which claimed many lives. Today, various developments have been made in farming using image processing, data mining, and machine learning techniques. The review's primary objective is to assist researchers in conducting pertinent machine-learning-based studies. This study highlights the most popular machine learning methods, including classification and regression trees and neural networks to detect the crop ailments. This assessment of the literature not only indicates that these strategies may compete with more established ones, but also shows how sensitive these techniques are to the training data.*
Keywords:*Machine Learning Techniques (MLT), Neural Networks (NN), crop ailments detection, image processing, agriculture.*

1. INTRODUCTION

Environmental changes have led to the fungus and other illnesses that have affected crops. Hence, in order to increase productivity, farmers must address current issues and take appropriate action. By analysing the leaves of the crops, IoT, machine learning, and neural network techniques are employed to identify crop illnesses. Studies is being done to identify the climate change sources that will have an impact on production and the illness detection training model that will be used. [2] The appropriate insecticides should be applied to the plant to prevent the spread of disease and boost yield.

To avoid wasting resources, it is important to conduct a thorough investigation of the plant diseases. By detecting changes in the plant's leaf and appropriately setting the system to identify the diseases, early disease detection can be made [1] [2] [4]. By using neural networks, the prediction is to be made in order to identify the disorders [3] [5]. Using a decision tree, the temperature and soil moisture sensor data from the prior and current years must be taken into account to provide adequate water management and prevent disease infection in the plant [1].

The model was trained to identify the illnesses affected in the plant using machine learning techniques [2] [4]. In order to construct a model, raw data on the diseases affecting the leaves are combined with a vast amount of data to identify the affected leaves [3] [5]. When trained data anduntrained data are compared, information regarding diseases can be given to the farmer to increase output.

This study focuses on several machine learning paradigms. The comparative analysis of machine learning techniques with appropriate application domains is also examined in our research. This article's remaining sections are organised as follows: This article's Section 2 discusses the literature review. Section 3 presents the various crop diseases, including Powdery mildew, Alternaria, and Grey Mildew for wheat and cotton crops. The major machine learning techniques, including SVM, K-NN, and CNN, are highlighted in section 4. The conclusion is presented in section5.

2. LITERATURE SURVEY ON LEAF DISEASES

Jayraj Chopda, et al [1] detection of the cotton crop diseases using decision tree algorithm with real time data from the sensor data like temperatures, sensors etc. are compared with the previous data and based on the obtained values the diseases like anthracnose, areolate and wilt are detected using decision tree algorithm. The decision tree classifier predicts the infected leaf with the temperature and the sensor data by the decision

prediction and informs the user about infected leaf.

Anuradha Badage[2] crop diseases is detected in the early stage by using edge detection and histogram matching. The threshold values are obtained for both aging and infected leaves based on the previous data is analysed with present values of the healthy and diseased leaves. With the obtained values the diseases are detected by canny edge detection.

Pooja Pawar et al [3] cucumber diseases are identified by the Artificial Neural Network by using image acquisition, image pre-processing, feature extraction of the leaf of cucumber this data is compared with trained dataset. The GLCM and statistical data is used to extract the texture feature and this system provides the classification accuracy of 80.45%

Shima Ramesh, Et al[4] plant diseases is detected by random forest algorithm accurate compared to the other technique SVM, Gaussian Navie Byes method the random forest is more accuracy with less number image dataset. The accuracy of the random forest classifier is approximately 80 percent.

Ritesh Sharma, et al [5] Neural Network is used to extract image features from the leaves to detect what type of disease the paddy plant is suffering. The images are augmented toincrease the training data. CNN has predicted dataset values which verified with actual values. CNN technique requires a large dataset to detect disease in the crop.

Bhumika S, et al [6] a survey on cotton diseases uses SVM to detect and classify various crop diseases image obtained from digital cameras. Pre-processing is applied in order to remove the background images. Segmented images are extracted to get the colour and shape of the image.

P. Revathi, et al [7] cotton leaf disease detected by using neural network. The trained dataset is used to identify the early detection of diseases. The image segmentation is done to get the exact spot of disease. The image boundary, shape and color are identified to find the diseases and the pest recommendation is done by supervised classification.

Yuan Tian,Chunjiang Zhao et al [8] the leaf diseases of wheat are identified by SVM, stacked generation structure and mid-level future generation is used to improve the performance to identify the diseases.

Sunil S. Harakannanavar et al [9] the computer vision techniques like RGB conversion to gray, HE, K-means clustering, contour tracing is equipped in pre-processing stage. Discrete Wavelet Transform, Principal Component Analysis and GLCM are used to extract the informative features of the leaf samples. The machine learning technique is implemented to detect diseased leaves from the healthy leaves.

3. CROP DISEASES

Diseasesinwheatcrop:

Thediseasesthatharmwheatcropsincludeleafrust,stem rust, powdery mildew, and yellow rust as illustrated infig.1.Theleafrustwillspreadfromthebaseoftheleaftothestem node and the leaf rust is normally small, round, andcauseslessleaftissue[2].Powderymildewusuallyappearsascircular, powdery white spots, which can appear on leaves,stems,andsometimesfruit.Itiscausedbynutrientdeficiencies, pests, and disease. Powdery mildew is caused bywhite lesions on leaves [2]. Yellow rust, or stripe rust, takesitsnamefromtheappearanceofyellow-coloredstripesproducedparallelalongthevenationsofeachleafblade[2].

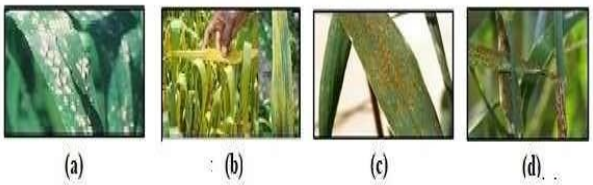


FIGURE 1. (a) Powdery mildew (b) Stripe rust (c) Leaf rust(d)Stemrust

Diseasesincottoncrop:

Xanthomonas citripv. malvacearum is the culpritbehind cotton's bacterial blight, also known as seedling blight,angular leaf spot, and boll rot. The bacterium can infect stems,leaves,bracts,andballsofthe cottonplant atanystageofitsgrowth. It results in boll rot, seedling blight, leaf spot, blackarm,andblackvein.Thesebacteriaontheleavesareareolate,measuring 1-2 mm on the bottom surface, and are dispersed insmall, dark-green, wet areas [2]. Brassica crops

frequentlycontractAlternarialeafspot,afoliardiseasebroughtonbythefungus Alternaria brassicicola.The disease is more likely todevelop in plants that have prolonged leaf wetness periods.Water encourages disease growth since viruses don't start togerminate and infect leaves until they've been exposed to it formorethanninehours[2]. ThefungusPseudocercosporapurpurea is the cause of cerea. Circular to asymmetrical leafpatches up to 10 mm in diameter were more or less vein-limited. When heavily fructified, the spots changed fromuniformlybrowntoreddishbrowntopalebrownwithapurple border and revealing patches of greyish skin on thelesion [2]. The fungus Botrytis cinerea is the source of theillnessknown asgrey mildew,which results in agentledegradationofplanttissuesandtheformationoffuzzy,grey-brownmold.Numerousplantsareimpacted,especiallythoseraisedinhumid,glass-coveredenvironments[2].Theyellowing, stunting, and death of seedlings as well as theyellowing and stunting of older plants are indications offusarium wilt. The lowest leaves of infected plants quicklyturn yellow and dry, and the xylem tissues turn brown and alsoit causes the plant to die [2].Fig.2 identifies the ailmentsthatinfluencecottonleaf.

FIGURE 2. (a) Alternaria (b) Bacterial blight (c)Bacterial blight (d)Grey Mildew (e) Fusarium Wilt

Frequentlyemployedmethodforfindingdiseasedleaves:
The approach, which uses a random forest classifier toclassify more leaf pictures, can accurately identify diseasedand infected leaves [4]. By feeding previously recorded valuesintotheneuralnetworkandtrainingittoproducetherequiredoutput, the proper diseases can be identified [3] [5]. Photosthatcanberetrievedfromaerialorsatellitephotosareusedasinputbyatrainedmodelthatusesmachinelearningtoidentifyplantdiseasesandassessthehealthoftheplant.

4. MACHINE LEARNING TECHNIQUES

Programmingcomputersusingmachinelearningallowsthemtolearnfromthedata.Thecomputerisforcedtolearn whilebeingspecificallyprogrammedtoproduceacertain outcome. The model could be descriptive to learn fromthe data or predictive to make predictions about the future. Thesupervised learning, unsupervised learning, and reinforcementlearningarethethreecategoriesofmachinelearningalgorithms.

Unsupervisedlearning:
Unsupervisedlearningusesmachinelearningalgorithmstomakedecisionsbasedoninputdatafromdatasetswithoutlabelledanswers.Unsupervisedlearningmethodsdonotclassifyorcategorizeanobservation.Clusteranalysis,themostwidelyusedunsupervisedlearningtechnique, is used to uncover hidden patterns or grouping indatathroughexploratorydataanalysis.Themostsimilarobjectsremaininonegroupandhaveminimaltonosimilarities. Whenusingtheassociationtechnique,anassociation rule is a form of unsupervised learning that is usedto discover correlations between variables in a big database.It identifies the group of objects that appear collectively inthe collection. Reinforcement Learning is a feedback-basedMachinelearningtechniqueinwhichanagentlearnstobehave in an environment by performing the actions andseeing the results of actions. For each good action, the agentgets positive feedback, and for each bad action, the agent getsnegative feedback.In ReinforcementLearning, theagentlearns automatically using feedback without any labeleddata. The agent interacts with the environment and exploresit by itself. In reinforcement learning the agent has to improvetheperformancebygettingthe maximumpositiverewards.

Supervisedlearning:
It is a type of machine learning where computers are taughtwith carefully "labelled" practice data and then anticipateresults using that information. After the training phase iscomplete, the model is tested using test data, and then itpredicts the output. In supervised learning, the training datagiventothemachinesactasthesupervisorthattrainsthemtopredicttheoutputproperly.Regressionandclassificationaretwo examples of supervised machine learning. If there is acorrelationbetweentheinputandoutputvariables,regressionproceduresareapplied.LinearRegression,RegressionTrees,Non-Linear Regression, Bayesian LinearRegression, andPolynomialRegressionaresomeoftheregressiontechniquesthatfallwithinthe category ofsupervised learning.ThevariousclassificationalgorithmsincludeRandomForest,DecisionTrees,LogisticRegression,SupportVectorMachines,KNN,anddeeplearning.Fig.3depictstheprocessfor creating a prediction model using the trained and testdataset.

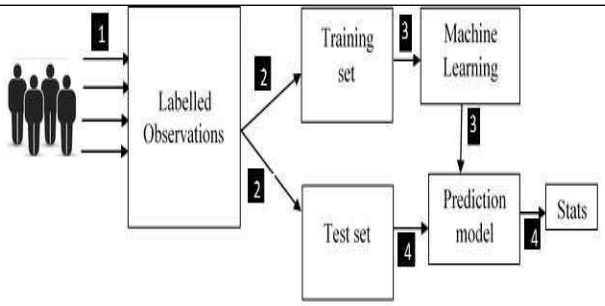


FIGURE 3. Supervised Learning

DecisionTreesalgorithm:

A decision tree is a supervised learning techniquethatcanbeusedtotackleclassificationandregressionissues,but it is frequently preferred for doing so. It is a tree-structuredclassifier, where internal nodes represent the features of adataset, branches the process of generating decisions, and eachleaf node the classification outcome. The creation of a leafnodegivenaclassificationoutcomeisshowninFig.4.According to the findings of their investigation, soil moisture,temperature, and a decision tree algorithm can all be utilizedtodetectillnessesincottoncrops.JayrajChopda,etal.[1]gottothisresult.

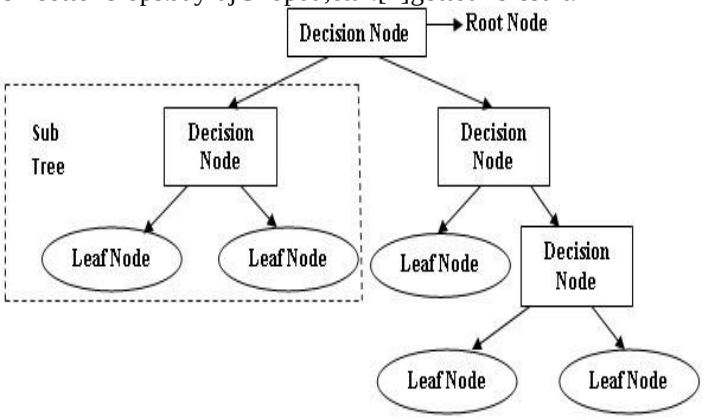


FIGURE 4. Decision Trees algorithm

RandomForestalgorithm:

A Random Forest classifier uses many decision treestocategorizedifferentsubsetsoftheinputdata.Toanticipatetheoutcomebasedonthevotesofthemajorityofprojections,the random forest uses predictions from all the trees ratherthanjustone.Themoretreestherearentheforest,themoreaccurate the results will be. The predictions that the randomforest algorithm produced for the data are shown in Fig.5.According to Shima Ramesh, et al. [4], the random forestmethodperformsbetterindetectingplantillnesseswithfewerphotodatasetsthanSVMandtheGaussianNavieByesmethod.

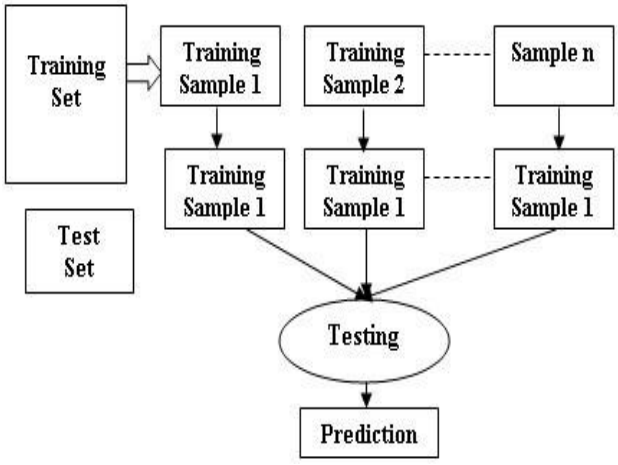


FIGURE 5. Random Forest algorithm

SupportVectorMachine:

Thetrainedmodel'sabilitytopredictthenewdataisshowninFig.6.TheSVMalgorithm'sobjectiveistoestablishtheb estlineordecisionboundarythatcandividen-dimensional space into classes, allowing us to quickly classifyfresh data points in the future. An SVM model, a supervisedmachinelearningapproachusedforclassificationandregression, may identify leaf illnesses by being given sets oflabeledtrainingdata [6],[8].Evenwithfewavailablesamples,SVMexecutesmoreefficientlyandquickly.

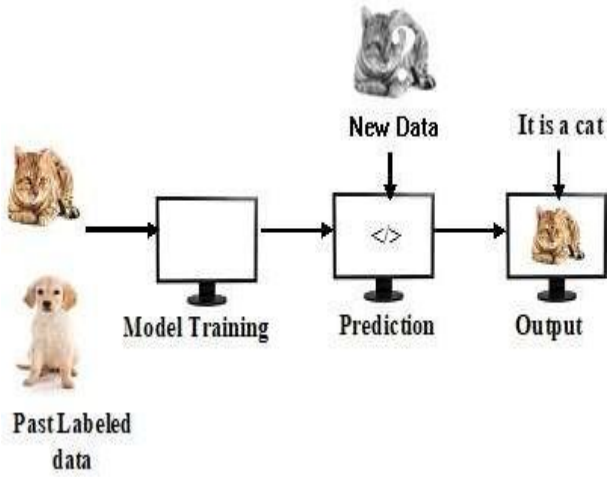


FIGURE 6. SVM algorithm in face detection

Deeplearning:

Deep learning models are computer files that datascientiststrainto performtasks using analgorithmorapredeterminedsetofsteps.Deeplearningmodelsareusedinvariousapplicationstoanalyzedataand generatepredictions.Deep neural network types include Multi-Layer Perceptrons(MLP),ConvolutionalNeuralNetworks(CNN),andRecurrentNeuralNetworks(RNN).Acompletel yconnectedmulti-layer neural network is referred to as a "MultilayerPerceptron" (MLP). Therearethreelevelstotal,oneofwhichisunderground.IfanANNhasmorethanone hiddenlayer,itisreferredtoasadeepANN.MLPsperformwellforclassificationpredictionproblemswheretheinput shavebeengivenaclassorlabel.Theyareeffectiveforproblemsinvolvingregressionprediction,whereareal-valuedquantityispredictedfromavarietyofinputs.Recurrentneuralnetworks (RNNs) are a form of artificial neural network thatareusedinspeechrecognitionandnaturallanguageprocessing(NLP).RNNisusedinbothdeeplearningandthe developmentofmodelsthatresemblethefiringofneuronsinthehumanbrain.RNNisfrequentlyusedinsentimentan alysis,textanalysis,machinetranslation,andimagecaptioning.ConvolutionalNeuralNetworks(CNN)andotherd eep learning algorithms are particularly good at processingand identifying images. By looking for patterns in the images,CNNs are highly useful for identifying items, classes, andcategories in photographs. When identifying the healthy anddamaged leaves in a paddy crop, the CNN algorithm canproduce a more accurate result [5]. Using image acquisition,imagepre-processing,andfeatureextractionofthecucumberleaf, Pooja Pawar et al. [3] claim that the artificial neuralnetwork can diagnose cucumber disorders. The leaf's health isthendeterminedbycomparingtheresultstotraineddatasets;fig.7showshowtheinputlayeristransformedintothe outputlayer.

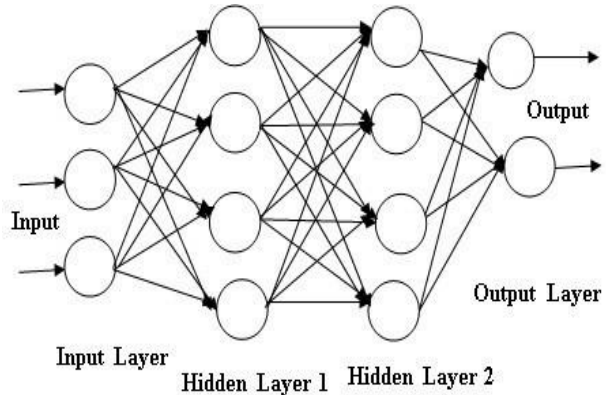


FIGURE 7. Deep learning model

K-nearestneighbour'salgorithm:

The k-nearest neighbour's algorithm, often known asKNNorK-NN,isamethodforsupervisedmachinelearning.As seen in fig.8, it is a non-parametric classifier that

usesproximity to make predictions or classifications about theclustering of a single data point. Assuming that the newscenario and previous ones are similar, the K-NN approachprovides a fresh instance in a group that appears closest andmost similar to the previous groups. K-NN does notimmediately learn from the training set; instead, it maintainsthe dataset and at the time of classification, it performs anaction on the dataset. KNN assumes that a value will beassignedtothenewdatapointbasedonhowmuchitresembles the points in the training set [9]. Table 1 providesresultsforvariousmachinelearningtechniques.

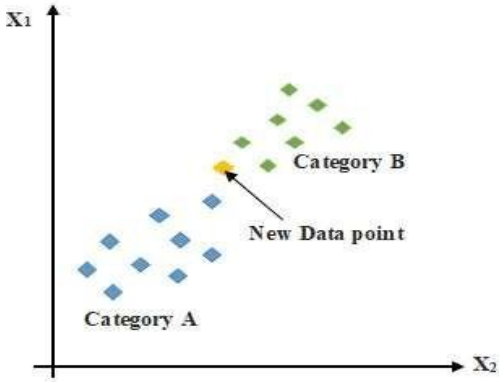


FIGURE 8. Identifying new data point using KNN

Comparisonofdifferentmachinelearningtechniques:

Byemployingmachinelearningtechniques,itispossible for computers to gain knowledge from data. Beingdeliberatelyprogrammedtoachieveaparticularresultforcesthe computer to learn. A descriptive or predictive model couldbe used to draw conclusions from the data and make futureprojections. The machine can identify the data more preciselyby using a classification method. The outcomes of differentmachinelearningapproachesaredisplayedinTable1.

Techniques used	Results	References	Accuracy
Decision tree algorithm	Cotton crop diseases is predicted by the previous data and real time data from the parameters like temperature, soil moisture	Jayraj Chopda, et al [1]	-
Canny edge detection	Crop diseases is identified by edge detection and histogram	Anuradha Badage[2]	-
Artificial Neural Network	Texture of image is detected by GLCM, statistical data and with neural network produce 80.4% accuracy	Pooja Pawar et al [3]	80. 45%
Random Forest Algorithm	Plant diseases is detected by random forest algorithm accurate compared to the other technique this algorithm is more accuracy with less number image dataset	Shima Ramesh, Et al [4]	70%
Convolutional Neural Network	This Neural Network is used to extract image features to detect what type of disease the plant is suffering. CNN has the predicted dataset values which verified with actual values.	Ritesh Sharma, et al [5]	90.32%
Support Vector Machine	Cotton and wheat diseases use SVM to detect and classify various crop diseases. The pre- processing and segmentation are done to get the colour, shape of the image. The stacked generation structure and mid-level future generation is used to improve the performance to identify the diseases	Bhumika S,et al[6], Yuan Tian,Chunjiang Zhao et al[8]	-
Segmentation, Edge detection	Cotton leaf disease is detected by boundary, shape, colour and texture are extracted for the disease spots using segmentation, edge detection.	P. Revathi, et al [7]	-
SVM, K-NN and CNN	The multiple descriptors Discrete Wavelet Transform, Principal Component Analysis and GLCM are used to extract the informative features of the leaf and SVM, K-NN and CNN used to identify infected leaf	Sunil S. Harakannanavar et al [9]	9 9%

5. CONCLUSION

The primary contribution of this article is a studyabout the different Machine-Learning Tools for analyzing leafphotosandtodetectcropdiseases.Thereportalsoprovidesacomparativeanalysisof allthestrategiesbasedontheiruses,benefits, and drawbacks. It is impossibleto identify onestrategy as the best after analysing every method. Based on itsstrengths, each strategy has a variety of applications and canbe effective across a range of fields. Therefore, the approachthat best fits a given application should be utilized, whilekeepinginmindboththelimitationsofeverymethodandtheoverarching objective of enhancing both productivity andperformance.Forexample,theRandomForestClassifieraidsin the identification of

diseased leaves, allowing farmers todeterminethetypeofdisease.Ourfindingsalsorevealedthatthere is no single machine learning technique that is perfect fordetecting crop disease and recommending pesticides.As a result, we have a specific goal to continue our researchonthesubjectbyrecommendingpesticidesforplantsthatareplaguedwithdiseaseandbyincorporating moredatatoidentifythetrueailments.

REFERENCES

[1]. Jayraj Chopda, Sagar Nakum,Vivek Nakrani, CottonCropDiseaseDetectionusingDecisionTreeClassifier
[2]. AnuradhaBadageCropDiseaseDetectionusingMachineLearning: IndianAgriculture
[3]. Pooja Pawar, Varsha Turkar, Pravin Patil, CucumberDiseaseDetectionUsingArtificialNeuralNetwork
[4]. ShimaRamesh,RamachandraHebb,NivedithaM,PoojaR, Prasad Bhat N, Shashank N, P V Vinod, Plant DiseaseDetectionUsingMachineLearning
[5]. Ritesh Sharma, Sujay Das, Mahendra Kumar Gourisaria,Siddharth Swarup Rautaray and Manjusha Pandey, A ModelforPredictionofPaddyCropDiseaseUsingCNN
[6]. BhumikaS.Prajapati, VipulK.Dabhi,HarshadkumarB.Prajapati, A Survey on Detection and Classification of CottonLeaf Diseases
[7]. P. Revathi, M. Hemalatha Advance computing enrichmentevaluation of cotton leaf spot disease detection using ImageEdgedetection
[8]. Yuan Tian,ChunjiangZhao,Shenglian Lu,SVM-basedMultiple Classifier System for Recognition of Wheat LeafDiseases
[9]. Sunil S. Harakannanavar,Jayashri M. Rudagi , Veena IPuranikmath, Ayesha Siddiqua, R Pramodhini,Plant leafdisease detection using computer vision and machine learningalgorithms
[10]. ChaitaliGDhaware,KHWanjale,Amodernapproachforplantleafdiseaseprocessingclassificationwhichdependso nleafimage
[11]. V Gupta , N Sengar , M Dutta , C Travieso , J Alonso ,Automated segmentation of pow- dery mildew disease fromcherryleaves usingimageprocessing
[12]. M Arya , K Anjali , D Unni , Detection of unhealthy plantleaves Using image processing and genetic algorithm withArduino
[13]. https://www.kisansuvidha.com/wp-content/uploads/2017/07/grey-mildew-of-cotton.jpg
[14]. http://www.hobbytronics.co.uk/arduino-uno-r3
[15]. https://cdn-hop.adafruit.com/product-files/2471/0A-ESP8266DatasheetENv4.3.pdf
[16]. https://www.itead.cc/wiki/DHT11HumidityTemperatureSensorBrickSpecifications
[17]. https://www.dfrobot.com/wiki/index.php/DHT11TemperatureandHumiditySensor(SKU:-DFR0067)
[18]. https://www.amazon.in/KitsGuru-Moisture- Sensor-ModuleArduino/dp/B00XU8MJ4E/ref=sr11?ie=UTF8qid=1509719422sr8keywords=soil+moisture+sensor
[19]. https://www.crowdflower.com/what-is-training-data/
[20]. https://thingspeak.com/