D-5893
[1-12]

REVIEW ARTICLE Agricultural Science Digest

Machine Learning Models for Plant Disease Prediction and Detection: A Review

Shivappa M. Metagar[1], Gyanappa A. Walikar[2]

10.18805/ag.D-5893

ABSTRACT

In the agriculture field farmers are dependent on crops and once it is caused by disease, they can lose the proper yield. Plant illnesses are the common cause of low yields and reduced income for farmers. So, to overcome this, machine learning approach in the agriculture field is showing exponential improvement for interdisciplinary research, but in modern disease prediction systems, it is a different process that can takes more time to identify and understand the type of disease. Disease prediction is related to computer vision and machine learning to detect types of leaf disease on different plants. It's a long-term system that includes a variety of real-world scenarios such as, Prediction system for vine leaf disease, pomegranate leaf disease, corn leaf disease, *etc.* Disease prediction is the accurate determination of the disease state of plant leaves. Accurate disease prediction is one of the key requirements in data science. Presently, machine learning based methods have improved prediction accuracy for plant leafs like grapes, pomegranates, maize. However, the disease prediction performance is still required to be improved in this challenging environment. Existing disease prediction models require high computational time and storage facilities in the agriculture field. To overcome this, we have proposed a comparative study of ML for prediction and classification of crop diseases to improve the efficiency of early prediction of crop diseases.

Key words: Classification algorithms, Decision tree, Feature extraction, K-nearest neighbor (KNN), Random forest.

Modern disease prediction systems are another process that can take a long time to understand the type of disease. Disease prediction is related to computer vision and machine learning and involves detecting the type of disease in different plants. This is a long-term research problem with many real-world applications, like pomegranate leaf disease, grape leaf disease *etc*. Disease prediction is related to finding the accurate disease of the plant. Accurate disease prediction is one of the crucial requirements in the field of data science. The performance of existing methods is not improved for disease prediction in this challenging environment. A prediction scheme is proposed, an ML model that improves the efficiency of disease prediction. With the advancement of new advancements, the area of study of agriculture becomes more important as it is also utilized as a meal for a multitude of the population but additionally in a few of additional uses. Floraiscrucial in our lives because they furnish a source of power and help to surmount the issue of climate change. Today, flora is impacted by a few diseases, as an example those that cause harsh economic, social and ecological damage, amongst other people. Therefore, it is of great importance to accurately and rapidly detect plant diseases.

The authors (Amrita and Rahul, 2019) have explained the visual observations are often used to determine the severity of disease in production areas. Image processing has made the greatest progress in agriculture. Diverse neural links approaches as an example back propagation and considerable component examination (PCA) have been utilized to recognize fungal diseases. Discover plant leaf diseases by enhancing the need rate in classification

[1]School of Computer Science and Engineering, REVA University, Bengaluru-560 064, Karnataka, India.
[2]School of Electronics and Communication Engineering, REVA University, Bengaluru-560 064, Karnataka, India.

Corresponding Author: Gyanappa A. Walikar, School of Electronics and Communication Engineering, REVA University, Bengaluru-560 064, Karnataka, India.

Email: gyanapp@rediffmail.com

How to cite this article: Metagar, S.M. and Walikar, G.A. (2023). Machine Learning Models for Plant Disease Prediction and Detection: A Review. Agricultural Science Digest. DOI: 10.18805/ag.D-5893.

Submitted: 04-10-2023 **Accepted:** 13-12-2023 **Online:** 26-01-2024

approaches. So far, responsible lenders can be found in this place. A linear SVM is utilized and this is a multi-class classification that categorizes the information into only two classes, but it has been unproductive and outcomes in less precise classification. Pests result in the destruction of crops or plant parts, lowering meal fabrication and guiding to meal insecurity.

Below equations and methods are used to find the type of the disease and accuracy percentage of disease in the plant and machine learning algorithms are used to find the type of the disease using classification algorithms and predicting the percentage of the disease attacked.

Evaluation models

Various types of measures are introduced to check the performance of vision-based and classification techniques. Accuracy, precision, recall and F1 scores are the most used checking measures for classification models. Parameters such as precession, recall and F (8), true positive (TP), false positive (FP), false negative (FN) and true negative (TN) for the proposed model the authors (Thakur *et al*., 2022) were calculated.

Accuracy

It's defined as the ratio of the complete number of right predictions to the complete number of samples (T). Precision is utilized interchangeably with recognition rate (RR), hit rate (SR) or precision rate (CR) in the introduction.

$$\text{Acc} = \frac{(\text{TP} + \text{TN})}{\text{T}} \quad ...(1)$$

Precision

Precision is defined as the percentage of accurate positives discovered over the complete number of predicted determined samples.

$$\text{Pre} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \quad ...(2)$$

Recall

Recall is defined as the ratio of correctly predicted positive outcomes to actual positive outcomes.

$$\text{Re} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad ...(3)$$

F1-score

The harmonic mean of precision and recall is defined as the F1-score.

$$\text{F1} = \frac{(2 * \text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \quad ...(4)$$

Specificity

Specificity is defined as the ratio of the exact pre-detected negative to the actual negative.

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN} + \text{FP})} \quad ...(5)$$

mAP

The intended precision was defined as the place below the accuracy-retract arc(AP). The intend ordinary exactness (mAP) is the ordinary of the intend precision AP_i, $i \in 1$......, n Over a batch of n samples.

$$\text{mAP} = 1/n \, \Sigma_{n=1}^{n}(\text{APi}) \quad ...(6)$$

This paper is structured according to the following manner. In section I, Brief introduction about plant leaf disease detection concepts, In the II section related work is outlined in detail, like supervised and unsupervised machine learning and in supervised machine learning again first section is regression algorithm, decision tree, random forest, classification algorithms like KNN, Naive Bayas, SVM and regression algorithm, in section III brief summary of all the machine learning algorithms, In section IV final conclusion is set.

Types of Machine Learning and its algorithms are used to classify and detect plant disease.

In Fig 1, the various types of supervised and unsupervised machine learning algorithms are listed that are used to find the different types of diseases and also classifying various types of the disease in plant.

- Traditional methods are the manual checking methods and takes more time and the plant disease detection made based on symptoms of the plant.
- The use of ML methods such as SVM, K-NN and Naive Bayas is used to separate diseased and uninfected leaves.
- Currently, machine learning-based methods have improved the accuracy of predictions for crops such as grapes, pomegranates, corn; however, disease prediction performance still needs to be improved in a harsh environment.

In the below line have explained about the different machine learning algorithms and uses of algorithms for implementation of various tasks in the agriculture field for identifying the various types of disease and percentage of the disease attacked to the plant.

Supervised machine learning

Well trained machines are used for supervised machine learning "labeled" training dataset based on this training dataset output will be predicted by machines. Labeled data means that some input data has been labeled with the correct output in the trained dataset. In supervised machine learning, training data is fed to a machine that functions as a supervisor that can teach the machine to correctly predict output given a given set of input data. It applies to the same concepts that students learn under the supervision of their homeroom teachers. It's a procedure that feeds the accurate input along with the result information to the artificial intelligence algorithm. The goal of a supervised acquiring procedure is to locate a mapping operate that maps the input factor (X) to the product factor (Y).

Regression

The equation used for linear regression can be written as:

$$Y = Mx+C \quad \text{OR} \quad y=bo+b*x+e$$

Where,

M and C = Gradient and constant values (to be determined) defining the straight line.

bo = Intecept,

b = Regression weight or coefficient associated with the predictor variable x.

e = Residual error.

It's a technique for locating the relationship between autonomous factors or characteristics and a measured factor or result and it's utilized as a way for predictive modeling in device learning, whereat a procedure is utilized to forecast continual result.

Investigating machine learning regression techniques for leaf rust detection using ultrasound measurement in this paper the authors (Davoud *et al*., 2016) have discussed the composite impact of symptoms and disease stages on the main characteristics of the plant leading to limiting the severity of the disease detected by different techniques. In addition, ML techniques are used to detect plant diseases; the influence of these symptoms on their performance is not taken into account. Spectra of infected and uninfected leaves under various disease symptoms were measured with a non-imaging spectrophotometer in the electrical region of 350-2500 nm.

The authors (Adhao *et al*., 2017). Have explained about the cotton is the most important cash crops in India and its influence on the economy. Cotton production is decreasing every year due to various disease attacks. Usually, crop diseases are caused by harmful insects and pathogens, reducing yields on a large scale if not promptly prevented. This work presents a system for the detection and control of cotton leaf disease.

Logistic regression

Calculating the probability, we used logistic regression to find the binary outcome: to check if something happened or not can be shown in Yes/No or Pass/Fail. Unbiased variables are analyzed to decide on a binary outcome, with consequences falling into one of two categories. The independent variables can be expressions or numbers, but the dependent variable is always an expression. Write like:

$$P = \frac{Y = 1}{X} \quad \text{OR} \quad P = \frac{Y = 0}{X}$$

Calculating probability is X given independent variable, the Y dependent variable.

The convolutional neural network using optimized proposed logistic decision regression (Priyanka *et al*., 2022), have discussed the agricultural nature is very important to cultivate crops with the aid of machine learning and artificial intelligence. In this work, we will detect leaf diseases by implementing image classification and analysis algorithms. Conventional manual identification of medicinal plants is a quick process that requires the help of plant identification specialists to find plant diseases.

Decision tree

In machine learning, when we know the cost of each probability and the outcome it will produce, it can be calculated using the following formula:

Expected Value (EV) = (First possible outcome × Probability of outcome) + (Second possible outcome × Probability of outcome) - Cost

Steps used for making decision tree

Step-I: Take a set of rows list (dataset) which can get to make for making decision tree.

Step-II: Dataset calculation for uncertainty or impurity in dataset or checking the mixed dataset *etc*.

Step-III: List of all questions generation which need to be asked to that node.

It is perfect for classroom problems because it can schedule instructions to a specific level. It works like a diagram, isolating elements of information into two similar categories, from 'stem' to 'branch' to 'leaf' where the final categories look more similar. It creates categories within categories, in the spirit of natural classification with limited human supervision (Rajesh *et al*., 2020) have explained the detection and classification of foliar diseases according to the decision tree in, agricultural production depends on the economy.

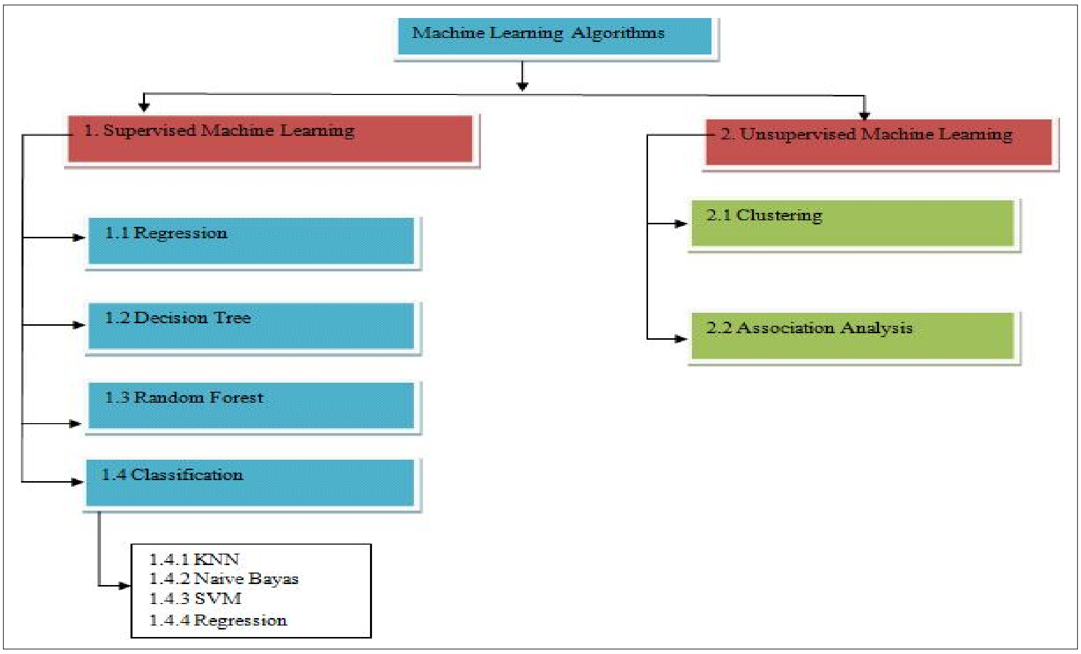

Fig 1: Main Structure of the machine learning types and its algorithms.

In Machine Learning domain the authors (Nalin *et al.*, 2023) have discussed the identification of plant disease by using different types of algorithms and they have been started for finding excellent result for data by using algorithm, this is an intriguing procedure in daily life application. The need for efficient algorithms is met when implementing these algorithms, the main objective of which is to focus on the prediction of plant diseases in the agricultural field in practice by providing resources in agricultural and business sectors. This bulletin's conceptual engineering inculcates tracking each leaf inside a tree through a machine learning model.

In Fig 2, the authors explained the gist that we use the results to provide, digest and approve a version used to predict fateful events and suggest possible effects. Capabilities while building predictive modeling methods. Regression and neural networks are two of the most well-known and widely used predictive modeling techniques. In a word, several types of Bayesian assessment, selection wood and statistical profile processing are complementary techniques. Each of the many prediction models has an important function to play. As we all understand, a metamodel is a model that can be used many times and is created by teaching an algorithm to use the information that has been collected, processed and saved for later use in the future evaluation of the final result.

Random forest algorithm

Random forest algorithm is a popular supervised machine learning technique; it is used to solve both classification and regression problems in machine learning.

Random forests converge

Given an ensemble of classifiers h1(x), h2(x) . . . hK (x) and with the training set drawn at. random from the distribution of the random vector Y, X, outline the margin characteristic as explained the in (Leo Breiman Statistics, 2001).

mg [(X, Y) = avkI(hk (X) = Y) - max j = Y avk I(hk (X) = j)].

In which I(•) is the indicator feature. The margin measures the volume to which the common wide variety of votes at X, Y for the proper magnificence exceeds the average vote for another elegance.

The extra self-belief within the category. The generalization errors are given *via*.

PE* = PX, Y [mg (X, Y) <zero]

Where in the subscripts X, Y indicate that the possibility is over the X, Y space.

A hybrid approach to detect and classify apple diseases using a random forest classifier in (Hernández *et al.*, 2020). Today, foreign trade has increased significantly in some countries. Range of overseas fruit products from trading countries, *e.g.*, oranges, apples, *etc.* Manual evidence to distinguish infected fruit is extremely tedious. The authors have (Meghana *et al.*, 2019) discussed the diagnosis of tomato plant diseases using the Random Forest algorithm, Random Decision Forest is a synthetic training method for regression, classification, *etc*. The random forest consists of a number of decision trees. The random forest summarizes the results of the entire decision tree at the time of training and class generation in the case of classification problems and average prediction in the case of regression model. Random decision trees overcome the problem of overfitting to their training set, which is the main problem of decision tree algorithms. In this work, they were divided into two phases. In the first stage, feature extraction is performed on all images and stored in the feature table.

In Fig 3, random forest decision tree algorithm are used to diagnose rice leaf diseases and includes several procedures like first obtaining an image of rice leaves; preprocess the images, extract capabilities from these images and classify the images by disease name and the main dataset is divided into many units. Two thirds of the records in the dataset, *i.e.*, a total of 352 images (276 captcha and the latest image added from) are available for training in the above concepts the authors (Sristy *et al.*, 2012).

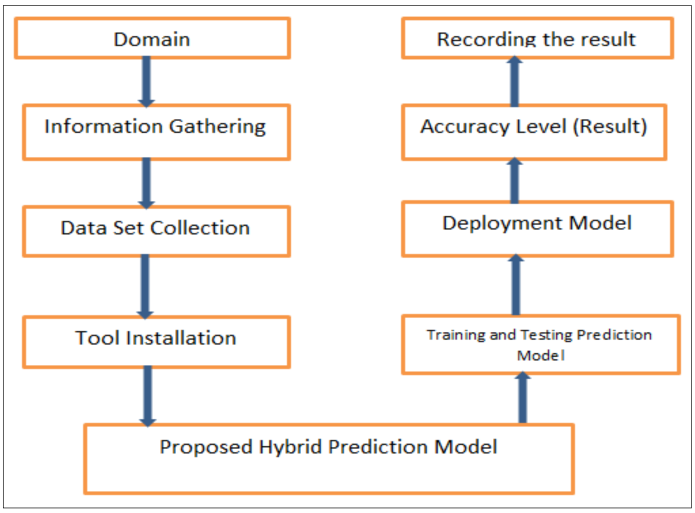

Fig 2: Domain based disease prediction and resource reference (Nalin *et al.*, 2023).

Classification algorithms

Class algorithms come under supervised device getting to know approach that is used to perceive the category of latest observations primarily based on the training dataset. In class set of rules an application learns the given dataset or observations after which classifies new remark into a number of groups.

K-nearest neighbor (KNN)

K-nearest Neighbor (k-NN) is a sample reputation set of rules that utilizes the education dataset to find k closest relatives in coming cases. Whereas k-NN is utilized in the layer, you work out to place the records in its closest neighbor kind. If OK = 1, then he could be placed in the lesson closest to 1. K is categorized by one vote as said by the number of his acquaintances.

In the Table 1, we have summarize the various KNN algorithms for finding and comparing the result of the different papers and in (Eftekhar *et al*., 2019) have explained the Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease Using KNN Classifier, Modern organic farming is gaining popularity in the agriculture of many developing countries. This work proposed a way to detect and classify leaf diseases using nearest neighbor classifier (KNN).

Naive bayes

Naive Bayes calculates the probability that a record point belongs to a particular category. Text analysis can be used to classify words and terms that belong to predefined tags (categories). The probability that X occurs if Y is true is equal to the probability that Y occurs if X is true multiplied by the probability that X is true divided by the probability that Y is true. Apply the formula for the probability of an event.

$$P\left(\frac{X}{Y}\right) = \frac{P\left(\frac{X}{Y}\right) * P(Y)}{P(Y)}$$

The authors have explained papaya plant disease (Wahyuni *et al*., 2020). Detection Using Fuzzy Naive Bayes Classifier Algorithm Papaya is one of the common fruits grown in Indonesia. With weather conditions infested with pests and diseases, a professional papaya disease detection system was developed. Expertise is reflected in the system, allowing farmers to determine status without an expert.

Support vector machine (SVM)

SVM uses algorithms that shape and classify data into levels of polarization to go beyond X/Y prediction to some extent. For simple image rationalization, we use tags: red and blue, the same two actual functions: X and Y and train a classifier to output X/Y coordinates to pink or blue. Rajleen and Kaur (2015) in Enhancement of SVM classification to improve plant disease detection, have proposed work in leaf imaging or even leaf imaging.

SVM errors = Margin error + classification error.

Here the higher the margin, the lower could-be margin errors and vice versa.

In this post, we run an SVM with records. One is school grades and train grades. First the original image is captured and then miles are used for processing. The image's black and legacy pixels are then segmented and the color and saturation portions of the image are also separated. Third, detect disease and sickness from photos and segment healthy parts from them. This study will tell you the percentage of the area where the disease occurs and may even give you the name of the disease. In Vagisha *et al.* (2020), the authors discussed leaf comparisons of different plants using SVM.

In the Table 2, have summarize the various SVM algorithm for finding and comparing the result, accuracy of the different author's paper. Machine learning techniques have been used by (Majji *et al*., 2021) to find plant diseases and traditionally, crop diseases are common causes of low yields and reduced yields. Accurate identification of plant diseases can help find cures as soon as possible to control losses. These authors attempted to develop a new method using ML techniques to predict plant diseases and compare different classification techniques; a comparison is given in Table 3.

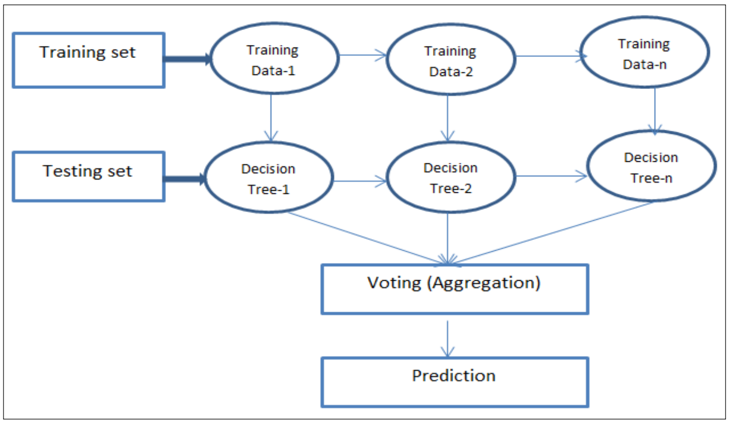

Fig 3: Random forest decision tree architecture.

Table 1: Work summary of KNN technique.

Title	Technique(s)	Plants used	Outcome/Accuracy
The detection and estimation of disease severity in cotton plants from unrestricted images are addressed in this studythe authors (Parikh *et al*., 2016) have explained.	KNN	Cotton plant	Precision of 82.5% was accomplished utilizing this scheme.
A comprehensive assessment for image-based detection of cassava disease (Ramcharan *et al*., 2017).	SVM and KNN	Cassava plant	An accuracy of 73% was achieved by implementing ANN and 91% by using SVM.
Recognition of diseases in rice Leaves utilizing KNN Classifier (Suresha *et al*., 2017).	KNN	Paddy plant	With this procedure the writers accomplished a basically precision of 76. 59%.
An advancement based on color and texture for the detection and classification of foliar diseases using the KNN classifier (Hossain *et al*., 2019).	KNN	Various plant leaves	The classification effectiveness of ANN in plant leaf diseases illustrates a precision of 96.76%. This development offers greater outcomes in comparison with several common methods.
A remote sensing technique for detecting avocado laurel disease in the presence of other biotic and abiotic assets (Abdulridha *et al*., 2019).	MLP and KNN	Avocado plant	In all cases, MLP approach acquired greater classification information than ANN, arriving 98% in several cases.

Table 2: Summary of work done using SVM technique.

Title	Technique(s)	Plants used	Outcome/Accuracy
The SVM classifier enables the detection of vine leaf diseases (Padol *et al*., 2017).	SVM	Grape plant	Treatment resulted in an average accuracy of 88.89 per cent for both downy leaf disease and Powdery grape leaf disease.
Detection of fungal rust in pea plants based on the SVM classifier analyze the (*Pisam sativam*) (Singh *et al*., 2019).	S	VM	Pea plant The submitted method can effectively identify and disease with an accuracy of 89.60%.
Disease detection in cotton leaf maps utilizing illustration processing and SVM classifiers (Bhimte *et al*., 2018).	SVM	Cotton plant	The offered design discovered diseases in cotton flora with 98.46% precision.
Identifying Plant Leaf Diseases utilizing Spider Monkey Index Optimization (Kumar *et al*., 2018).	SVM	Potato and Apple plant	Precision of 92.12% has been acquired utilizing this procedure.
Tea leaf disease detection and detection using Assist Vector devices (Hossain *et al*., 2018).	SVM	Tea plant	The method accomplished generally speaking precision of 93.33%.
Banana plant disease detection and classification utilizing local binary sample and assist vector device (Aruraj *et al*. , 2019)	SVM	Banana plant	The highest precision that the offered work attained was 89.1% and 90. 9%.

Table 3: Comparison of classification techniques with strengths and weaknesses.

MLalgorithms	Strengths/Potency	Weaknesses/Exhaustion
SVM	It has excellent capabilities for weed identification.	1) It needs longer education and testing if the effectiveness of the segmentation is sluggish. 2) It's not proper for big datasets.
KNN	No prior expertise in education is demanded. This reduces detection time and procedure complexity.	Choosing a k-value is challenging.
NB	Basic classifier, elevated accuracy and sprint with big database.	It works on offline information only
BPNN	Elevated precision.	More education time is demanded
DT	Elevated precision.	Over adapting problem has to have. More education time.
RF	Elevated accuracy, handles absent values, violates over adapting issue.	It needs more algorithmic electrical power and education time.

Table 4: Machine learning algorithms comparison.

Machine learning technique	Benefits for using ML models	Losses for using ML models
Decision tree learning	1) Aptitude to choosing the most discriminatory characteristics. 2) Handling both continual and discrete information.	1) Actually, they are unstable. 2) And also, calculations can get complicated.
Association rule learning	1) Finds similar patterns in data and produces rules.	1) Utilizes acquisition, integration and integrity check methods.
Support vector machine	1) Supervised acquiring technique. 2) It helps for classification and prediction aims.	1) Creates extremely precise classifiers. 2) Less overfitting and handling racket.
Clustering technique	1) Find commonplace grouping. 2) To generatean educated guess from information. 3) Discover dependable. Institution of information.	1) Offers an end user with elevated level display of what is transpiring database. 2) Extremely effective technique.
K-nearest neighbor	1) Strong too loudly education information. 2) Persuasive if education information is big.	1) Need to choose the importance of the parameter of k. 2) The computation price is elevated.
Genetic algorithm	1) Computer programming genetic algorithmic program is simple.	1) it's difficult for individuals to think of excellent heuristic. 2) Perhaps not discover an optimal approach to defined issue in all cases.

Table 5: Summary of different machine learning algorithms.

ML Approach and authors	Model design	Highlight	Research gap/Lacuna
ML and IoT Model by authors Pawara *et al*., (2019).	Timely identification of pomegranate disease utilizing artificial intelligence and web of belongings.	1. Quality and number of agricultural cosmetics are decreasing. 2. Low generator and obtaining low earnings. 3. Climate and Humidity sensor DHT22 will forward readings of heat and humidity to arduino that will actasand user interface between sensor and GSM module.	1. Because of not following the appropriate in the moment detection of disease besides particular use of pesticide. 2. As a consequence of the environmental factors, pests, diseases and not spraying of chemical pesticides. 3. Input Feed to treat is information gathered from all detectors. The method contains an artificial intelligence procedure trained on current information.
ML Model by authors Gawande Apeksha *et al*. (2021).	A concise survey on the prediction of chop disease utilizing. artificial intelligence approaches.	Disease and their respective symptoms can be used to support the early diagnosis of rice disease through Bayesian network with emphasis on naïve bayes.	1. Loss of crops from plant diseases is a rapidly rising contest globally and in the approaching 100 years. 2. We must avert starvation and starvation with several stronger tools and precise ML methods.
DL Model by authors Asta Lakshmi *et al*. (2021).	Automatic prediction of plant leaf diseases utilizing profound acquiring versions: A review.	1. Advancement of automated era in agriculture and meal manufacturing, there're however a few phases that need enhancement. 2. Diverse diseases like blights, leaf spots other bacterial and fungal infections are found in plant leaves.	1. As said by the effects, a transfer mastering machine with greater generally speaking efficiency is to be conducted alongside the Jetson Nano package for quicker and effective processing. 2. ML classification algorithm can be used to find the defects in the quality and quantity of the plant.
CNN binary classifier by authors Ahmad and Singh (2021).	Comparative survey of disease detection in flora utilizing artificial intelligence and profound acquiring.	Apple and tomato Leaves are used to detect disease.	Less accuracy because less number of data set used.
SVM, Random Forest, AlexNet, Google net same authors in 2021.	Comparative survey of disease detection in flora utilizing artificial intelligence.	SVM algorithm is 94.38%, Random forest algorithm is 95.46%, Alex N*et al*gorithm is.98.6% GoogleNet is 99.185% respectively.	Test accuracy on harsh condition with good result because of more data set and different algorithms.
ML Model by authors Radwan Qasrawi *et al*. (2021).	Machine learning techniques for tomato plant diseases clustering, predictionand classification	1. Six versions of ML (Neural networks, logistic regression, random forest, SVM, decision trees and Naive Bayes) were used in this study. 2. All versions used image-based datasets with real positive results. Optimal recognition accuracy and classification efficiency.	1. Disease clustering model, clustering with 70% accuracy. 2. Other models have used neural connections and logistic regression models but reported efficiency accuracies of 70.3% and 68%. 9% or.
The authors Raj Kumar and Neha Shukla, Princee have prepared the CNN and ML Model in 2022.	Plant disease detection and crop recommend and on using CNN and machine learning.	1. Crop and plant disease prediction plays an important role in agriculture. 2. An automated scheme to remediate plant diseases based on plant symptoms could be useful and beneficial.	1. ML models will be extremely helpful for resolving the noted issue. 2. The most important aim of this put forward scheme is to find plant diseases utilizing illustration processing.

Table 5: Continue...

Table 5: Continue...

The authours Hareem Kibriya *et al*. Have prepared the CNN and SVM in 2021.	Plant disease identification and classification using convolutional neural network and SVM.	1. The economy of agricultural countries depends mainly upon agricultural production. 2. Crops might be impacted by diseases because of diverse influences as an example climate change, pests, *etc*. that is able to harm the crops heavily.	1. Current methods of identifying plant diseases are tedious and error-prone because they require artisanal option extraction and segmentation. 2. For this reason, powerful automated methods developed by (CNN) are used in certain research areas.
ML and DL techniques prepared by Lakshmanarao, N. Supriya *et al*. (2022).	Plant disease prediction using transfer learning techniques.	1. Discovering and detecting plant disease is crucial in agricultural fabrication. It involves a lot of time and attempt to locate the disease. 2. Agricultural zone can additionally reap the advantages of ML and DL. There was a newest upward propel inside the application of ML and DL techniques in plant disease identification.	1. The prepared three datasets have been utilized for testing. All the experiments are conducted with python. Programming language. 2. We utilized a 'plant village' dataset gathered from Kaggle to give the correct consequence.
ML model prepared by Ruthvik Kilaru and Kommisetti (2021).	Prediction of maize leaf disease detection to enhance chop generate utilizing artificial intelligence founded versions	1. Corn husk production is affected by plant leaf diseases. 2. Maize is one of the most important crops among crops and is the main source of energy for human food and has achieved efficiency gains across 3. Farmers face the challenge of regulating and identifying crop diseases that affect the quantity and quality of maize crops in upland fields.	1. Early detection of the disease is critical to avoiding this large loss and increasing the productivity of the corn crop. In this study, five of his ML methods are used in maize India leaf disease detection: SVM, RF, ANN, DT and NB. 2. Input information is taken from the kaggle plant Village dataset and these annotated images are used to form a classification algorithm to find diseases.
Ensemble methods are prepared by Chandraprabha *et al*. (2021).	Prediction of crop diseases to improve crop yield using ensemble methods.	1. Research connected with agriculture is increasing swiftly; the insisting on situations that lie beforehand is solved with the aid of growth in era. 2. To improve the development and fix the issues mining methods.	Ensemble methods have been utilized to predict the chop disease and an analyze has been carried out with the aid of distinct classifiers besides option bushes, neural networks, naive in the agricultural sector, distinct information bayes classifier, random forests, aid vector method and okay-closest Neighbor.
Prediction using SVM explained by Farhana Tazmim Pinki *et al*. (2021).	Content based paddy leaf disease recognition and remedy prediction using support vector machine.	1. It is far very time-ingesting and onerous for farmers of remote regions to pick out paddy leaf sicknesses due to unavailability of professionals.	1. Experts are to be had in a few regions; sickness detection is completed *via* naked eye which reasons irrelevant recognition sometimes. 2. The write of rice leaf diseases has been accredited by assist Vector device (SVM) classifier.

The authors (Ramesh *et al*., 2020) have explained machine learning methods for the detection and classification of foliar diseases, a detailed review of the benchmarks of various advanced ML algorithms for identification. Identification and classification of foliar diseases.

Comparison of machine learning techniques

Strengths and weaknesses of various machine learning techniques are discussed in Table 3 and 4.

In the Table 5, we have explained the comparison between the different machine learning algorithms ML approaches and what they have designed the model, highlight of that model and gap in the research. Agriculture provides food for everyone, even in times of rapid population growth. Unfortunately, however, the disease occurs in the main stages of crops (Sunil S. Harakannanavar *et al*., 2022). The notion behind this document is to enlighten farmers on state-of-the-art to decrease leaf diseases in crops.

CONCLUSION

In this paper, with a broad instruction that covers detail concepts about the plant disease and summary of the plant disease classification and type of the plant disease, to suggest enhancement in current classification methods for plant disease detection using machine learning and that is verified by classifier procedure and an extra goal was to examine on distinct leaf diseases that are not yet been studied. This can motivate researchers to find new things and explore the use of numerous environmental factors which result in plant diseases which end up in giant production losses. It is a time-eating and blunder-prone system and it is able to be an unreliable method of identifying and stopping unfolding of plant diseases. Adopting superior technology such as Machine Learning (ML) can help to conquer those demanding situations by enabling early identification of plant diseases. A few diseases predictive studies have been conducted investigating recent advances in the use of ML techniques to identify diseases. The following are some directions to further improve the ability to detect plant diseases and increase work efficiency. In the future, we may do this work on more datasets to improve results and increase accuracy. This section outlines some of the research factors that need to be considered in future studies, followed by detailed conclusions.

ACKNOWLEDGEMENT

Authors acknowledge the support from REVA University for the facilities provided to write this article and thanks reviewers for their valuable suggestions.

Conflict of interest

The authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

REFERENCES

Abdulridha, J., Ehsani, R., Abd-Elrahman, A., Ampatzidis, Y. (2019). A remote sensing technique for detecting laurel wilt disease in avocado in presence of other biotic and abiotic stresses. Computers and Electronics in Agriculture. 156: 549-557. https://doi.org/10.1016/j.compag.2018.12.018.

Adhao A.S., Pawar, V.R. (2017). Machine learning regression technique for cotton leaf disease detection and controlling using IoT. International Conference on Electronics, Communication and Aerospace Technology. ICECA 2017, 978-1-5090-5686-6/17/$31.00 ©2017 IEEE.

Ahmad, N., Singh, S. (2021). Comparative Study of Disease Detection in Plants using Machine Learning and Deep Learning. Second International Conference on Secure Cyber Computing and Communication (ICSCCC), © IEEE.

Amrita, S., Raul, T.N. (2019). Plant Leaf Disease Detection using Machine Learning" 10th ICCCNT 2019, IIT - Kanpur, Kanpur, India.

Aruraj, A., Alex, A., Subathra, M.S.P., Sairamya, N.J., George, S.T., Eward, S.E.V. (2019). Detection and Classification of Diseases of Banana Plant Using Local Binary Pattern and Support Vector Machine, 2019 International Conference on Signal Processing and Communication (ICSPC -2019), March. 29-30, 2019, Coimbatore, India. 978-1-7281-1849-9 ©2019 IEEE.

Asta, L., Gomathi, M.V. (2021). Automatic Prediction of Plant Leaf Diseases Using Deep Learning Models: A Review. 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (ICEECCOT).

Bhimte, N.R. and Thool, V.R. (2018). Diseases Detection of Cotton Leaf Spot Using Image Processing and SVM Classifier. Second International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, pp. 340-344. DOI: 10.1109/ICCONS.2018.8662906.

Chandraprabha, M.T., Singh, A.S. (2021). Prediction of Crop Diseases to Improve Crop Yield Using Ensemble Methods. 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) | 978-1- 6654- 3811- 7/ 21/$ 31.00 2021 IEEE | DOI: 10.1109/ ICAC3N53548.2021.9725472.

Davoud, A., Aghighi, H., Matkan, A.A., Mobasheri, M.R. and Rad, A.M. (2016). An investigation into machine learning regression techniques for the leaf rust disease detection using hyperspectral measurement. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 9(9): 4344-4351, 2016 1939-1404 IEEE.

Eftekhar, H., Hossain, M.F. and Mohammad and Rahaman, A. (2019). A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease using KNN Classifier. International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February 2019. 978-1-5386-9111-3/19/$31.00 ©2019 IEEE.

Farhana, T., Khatun, P.N., Islam, S.M.M. (2021). Visual features based paddy leaf disease recognition, its severity detection and remedy prediction using k-means clustering and adaboost. Journal of Image Processing and Pattern Recognition Progress. 7(3): 41-52.

Gawande, A.R., Sherekar, S.S. (2021). A brief study on the prediction of crop disease using machine learning approaches. 978-1-6654-2040-2/21/$31.00 © IEEE.

Hareem, K., Abdullah, I., Nasrullah, A. (2021). Plant Disease Identification and Classification using Convolutional Neural Network and SVM. 2021 International Conference on Frontiers of Information Technology (FIT) | 978-1-6654-0830-1/21/ $31.00 ©IEEE |.

Hernández, S., López, J.L. (2020). Uncertainty quantification for plant disease detection using Bayesian deep learning. Elsevier Science Publishers B.V. Netherlands available on https://doi.org/10.1016/j.asoc.2020.1065971568-4946/© 2020 Elsevier.

Hossain, E., Hossain, M.F. and Rahaman, M.A. (2019). A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease Using KNN Classifier. 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox'sBazar, Bangladesh, 2019, pp. 1-6. DOI: 10.1109/ECACE.2019. 8679247.

Hossain, S., Mou, R.M., Hasan, M.M., Chakraborty, S. and Razzak, M.A. (2018). Recognition and detection of tea leaf's diseases using support vector machine. 2018 IEEE 14th International Colloquium on Signal Processing and its Applications (CSPA), Batu Feringghi, 2018, pp. 150-154. DOI: 10.1109/CSPA.2018.8368703.

Kumar V., Jat, H.S., Sharma, P.C., Singh, B., Gathala, M.K., Malik, R.K. *et al*. (2018). Can productivity and profitability be enhanced in intensively managed cereal systems while reducing the environmental footprint of production? Assessing sustainable intensification options in the breadbasket of India. Agriculture, Ecosystems and Environment. 252 (2018): 132-147.

Lakshmanarao, A., Supriya, N., Arulmurugan, A. (2022). Plant Disease Prediction using Transfer Learning Techniques. 2022 Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT) | 978-1-6654-1120-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICAECT54875.2022.9807956.

Leo Breiman Statistics Department, (2001). Random forests are a combination of tree predictors such that each tree depends, University of California, Berkeley, CA 94720, Machine Learning, 45, 5-32, 2001 c 2001 Kluwer Academic Publishers.

Majji, V., Kumaravelan, A.G. (2021). A Review of Machine Learning Approaches in Plant Leaf Disease Detection and Classification. Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV 2021). IEEE Xplore Part Number: CFP21ONG-ART; 978-0-7381-1183-4.

Meghana, G. and Veena, M. (2019). Diagnosis of Tomato Plant Diseases using Random Forest. 2019 Global Conference for Advancement in Technology (GCAT) Bangalore, India. Oct 18-20, 2019, 978-1-7281-3694-3/19/$31.00 ©2019 IEEE.

Nalin, K.Y., Maddali, K., Jayapandian, N. (2023). Machine Learning based Plant Disease Identification by using Hybrid Naïve Bayes with Decision Tree Algorithm. 5th International Conference on Smart Systems and Inventive Technology (ICSSIT 2023) 978-1-6654-7467-2/23/$31.00 ©2023 IEEE.

Padol, P.B. and Yadav, A.A. (2016). SVM classifier based grape leaf disease detection. Conference on Advances in Signal Processing (CASP). pp. 175-179. DOI: 10.1109/CASP. 2016.7746160.

Parikh, A., Raval, M.S., Parmar, C. and Chaudhary, S. (2016). Disease Detection and Severity Estimation in Cotton Plant from Unconstrained Images. 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Montreal, QC, 2016, pp. 594-601. DOI: 10.1109/ DSAA.2016.81.

Pawara, S., Nawale, D., Patil, K., Mahajan, R. (2018). Early Detection of Pomegranate Disease Using Machine Learning and Internet of Things. 3rd International Conference for Convergence in Technology (I2CT), India.

Priyanka, C., Gupta, S.S., Patnaik, D.M.S.P.K., Munagala, N.V.L.M.K., Sivasangari, A. and Tannad, H. (2022). Efficient plant disease prediction based on convolutional neural network using optimized proposed logistic decision regression. 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS). 978-1-6654-7657-7/22/$31.00 ©2022 IEEE.

Radwan, Q., Amro, M., Zaghal, R., Sawafteh, M. (2021). Machine Learning Techniques for Tomato Plant Diseases Clustering, Prediction and Classification. International Conference on Promising Electronic Technologies (ICPET) ©IEEE.

Raj, K., Shukla, N., Princee, (2022). Plant Disease Detection and Crop Recommendation Using CNN and Machine Learning. 2022 International Mobile and Embedded Technology Conference (MECON) 978-1-6654-2020-4/22/$31.00 © IEEE.

Rajesh, B., Vardhan, M.V.S., Sujihelen, L. (2020). Leaf Disease Detection and Classification by Decision Tree Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020) IEEE Xplore Part Number: CFP20J32-ART. ISBN: 978-1-7281-5518-0.

Rajleen, K. and Kang, S.S. (2015). An Enhancement in Classifier Support Vector Machine to Improve Plant Disease Detection. 2015 IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE). 978-1-4673-6747-9/15/$31.00 c 2015 IEEE.

Ramcharan, A., Baranowski, K., McCloskey, P., Ahmed, B., Legg, J., Hughes, D.P. (2017). Deep learning for image-based cassava disease detection. Frontiers in Plant Science. 8: 1852. https://doi.org/10.3389/fpls.2017.01852.

Ramesh, A., Schbert, R.D., Greenfield, A.L., Wilson, M.R. (2020). A pathogenic and clonally expanded B cell transcriptome in active multiple sclerosis, Edited by Lawrence Steinman, Stanford University School of Medicine, Stanford, CA and Approved August 3, 2020.

Ruthvik, K. and Kommisetti, (2021). Prediction of Maize Leaf Disease Detection to improve Crop Yield using Machine Learning based Models. 2021 4th International Conference on Recent Trends in Computer Science and Technology, (ICRTCST). DOI: 10.1109/ICRTCST54752.2022.9782023.

Singh, K., Kumar, S. and Kaur, P. (2019). Support vector machine classifier-based detection of fungal rust disease in pea plant (*Pisam sativam*). International Journal of Information Technology. 11: 485-492 (2019).https://doi.org/10.1007/s41870-018-0134-z.

Sristy, S., Ahsan, S.M.M. (2012). Rice disease detection using intensity moments and random forest. International Conference on Information and Communication Technology or Sustainable Development (ICICT4SD). 27-28 February, Dhaka, 978-1-6654-1460-9/21/$31.00 ©IEEE.

Sunil, S.H., Rudagi, J.M. (2022). Plant leaf disease detection using computer vision and machine learning algorithms. Global Transitions Proceedings, Journal Homepage. 3(11): https://doi.org/10.1016/j.gltp.2022.03.016.

Suresha, M., Shreekanth, K.N. and Thirumalesh, B.V. (2017). Recognition of diseases in paddy leaves using knn classifier. 2017 2nd International Conference for Convergence in Technology (I2CT), Mumbai, 2017, pp. 663-666. DOI: 10.1109/I2CT.2017.8226213.

Thakur, P.S., Khanna, P., Sheorey, T., Ojha, A. (2022). Trends in vision-based machine learning techniques for plant disease identification: A systematic review. https://doi.org/10.1016/j.eswa.2022.118117. Expert Systems with Applications 208 (2022) 118117Available online 11 July 2022 0957-4174/© 2022 / Elsevier.

Vagisha, S., Verma, A., Goel, N. (2020). Classification techniques for plant disease detection. Retrieval Number: F9902038620/2020©BEIESPDOI:10.35940/ijrte.F9902.038620.

Wahyuni, E.S., Kurniawati, Y.E. and Santosa, P.I. (2020). Papaya Disease Detection using Fuzzy Naïve Bayes Classifier 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI) | 978-1-7281-8406-7/20/$31.00 ©IEEE | DOI: 10.1109/ISRITI51436.2020.9315497.