

# **PLANT LEAF DISEASE DETECTION BY USING DIFFERENT CLASSIFICATION TECHNIQUES: COMPARATIVE**

---

## **ABSTRACT**

One of the main factors that assist to increase the growth of any country is Agriculture. The detection of diseases from plant leaf images is one of the most important fields of agricultural research. To identify disease factors, this field requires a reliable prediction approach. Data Mining (DM) is the process of analyzing data from different aspects and summarizing it into valuable information. It helps users to categorize and identify relationships between data from various dimensions. As there are many plants on the farm, detecting and classifying the diseases of each plant on the farm is extremely difficult for the human eye. And diagnosing each plant is very critical since these diseases may spread. DM classification is an important method that has a wide range of applications. It classifies each item in a set of data into one of a set of predefined classes. In this paper, a comparison of different DM classification methods such as Naive Bayes, Decision trees, SVM, and Random Forest algorithms has been illustrated by using of Weka Tool.

*Keywords: Data Mining, Plant Disease, Image Processing, Decision Tree, Naive Bayes, Random Forest.*

## **1. INTRODUCTION**

Data Mining (DM) is defined as a process that extracts valuable information from a huge data set. The majority of plant leaf disease detection systems have been designed in recent years using image processing and DM techniques [1] [2] [3]. Over the years, different algorithms have been produced, constructed, and modified to extract information from data sets. This also provides machine vision for automatic process control based on images [4] [5] [6]. Due to the high costs of chemical monitoring, epidemics allow more effective use of agrochemicals by detecting and differentiating several diseases in the early stages. The complicated agriculture problems seem to be solved by the DM methods, the process of collecting valuable and useful data for large data sets [7]. Agriculture's value is rapidly increasing in a country like India, where the population is so enormous, to satisfy the country's foodgrain self-sufficiency. As a result, to feed the entire world, it is also necessary to protect this from multiple threats and increase plant productivity [8] [9].

Plant disease detection which is done early is critical for maximizing crop yield. Plant diseases such as bacterial spots, black rot, black measles, and others influence plant growth, crop quality, and economics of the agricultural sector [10] [11] [12] [13]. Farmers also employ costly methods and chemicals to ward off the effects of these pathogens. The use of chemical means causes harm to the plant as well as the environment. Furthermore, this method raises the cost of production and causes farmers to lose a significant amount of money. The most critical timeframe for effective disease management is the early detection

of infections as they arise [14]. In agriculture, for assorting and identifying diseases of plants detecting diseases manually by specialists is a usual practice. Through computer vision and artificial intelligence research, automatic detection of plant diseases from raw images is now possible, due to technological advancements [15].

DM has many classification methods that assign items in a collection to target categories or classes. Classification aims to correctly predict the target class for each case in the data. A decision is used to classify data. Each decision is based on a query about one of the input variables. The data instance is categorized based on the acknowledgments. To obtain useful data, the classification method employs algorithms such as decision trees. Decision trees sometimes branch out to form random forests. These are classifiers with Meta estimators that can combine several decision trees into a single decision tree. The random forest's job is to increase the predictability of the controls and models used. Naive Bayes is a simple classification algorithm that predicts the classification of new data based on historical data. It measures the likelihood of an event occurring provided the occurrence of another event can be used for anomaly identification, regression, and classification. SVMs are based on the concept of determining the best hyper-plane for dividing a dataset into two groups.

The organization of this paper is as follows: Section 2 Related Work, Section 3 Classification Methods, Section 4 Plant Leaf Disease, Section 5 Methods, Section 6 Result and Discussion. Finally, in Section 7 the paper is ended with a conclusion.

## **2. RELATED WORK**

A comprehensive review of various machine learning algorithms for plant disease prediction is published, which will aid in the early identification of diseases and the implementation of preventative measures to preserve crop quality, minimize losses, and ultimately increase the yield of food crops[2]. In general, the appearance of the leaf indicates the presence of plant disease. If the leaf is determined to be an infected leaf, the resulting plant is determined to be diseased. To recognize diseased leaves, leaves are harvested from plants, and image processing is used to remove features. The model is then built using a machine learning algorithm, and leaves are evaluated against the model to determine the leaf is affected. The researchers paid close attention to disease detection in various plant species.

A method for detecting and categorizing diseases of plant leaf has been illustrated by Hossain et al. [16] by using a KNN classifier. From the images of leaf disease, texture features have been collected, for the classification. The KNN classifier would be used to classify diseases of different plant species such as bacterial blight, leaf spot, Alternaria, anthracnose, and canker. The proposed approach would successfully detect and classify the chosen diseases with 96.76% accuracy.

Kumar et al. [17] suggested a method for identifying leaf diseases by using different images. For collecting multiple features, they used segmentation techniques such as k-means clustering. To classify various kinds of diseases, the Gray Level Co-occurrence Matrix (GLCM) and the SVM classifier are used. The method aids in the precision detection of various diseases in leaves. The dataset is made up of multiple leaf images that have been affected by illnesses such as Cercospora leaf spot, bacterial blight, anthracnose, and Alternaria alternate. The findings are extremely accurate in identifying the affected region.

Kuricheti and Supriya [18] developed an algorithm for identifying and avoiding disease spread across the entire crop, resulting in crop production with high quality. The database of various leaf images was processed and developed utilizing k-Mean's image segmentation, and the analysis of leaf images was performed using GLCM. After rating the characteristics of the feature extracted images using an information gain algorithm, the SVM classifier is

used to classify them. A graphical user interface (GUI) has been designed to show the different stages of the image processing algorithm and to identify the two leaf diseases.

Bin Abdul Wahab et al. [19] image-processing algorithm based on Artificial Intelligence is proposed to identify diseases on a Chili plant utilizing images of its leaves. The proposed solution compares various SVM algorithms for classification and uses the k-means clustering algorithm for image segmentation. For classifying the images the features of images that are computed have been collected and used. To calculate different SVM classification algorithms, different parameters and kernel functions are used. The findings are divided into three categories: background, uninfected, and infected (Cucumber Mosaic), and they can differentiate between uninfected and infected plants.

Rajesh et al. [20] suggested a system for detecting plant leaf disease using a decision tree algorithm. The proposed method provides details on classification strategies for disease detection in plant leaves, and also algorithms for image partitioning methods for disease detection and classification in leaves. Early identifying infections of plants are also an advantage of this method. When it comes to detection accuracy, the planned work outperforms the current system.

Militante et al. [21] suggested a simple method for identifying various diseases in a variety of plant species. Apple, rice, grapes, potato, sugarcane, and tomato were among the plant species that the system was intended to identify and accept. The system can also detect a variety of plant diseases. The trained model had a 96.5% accuracy rate, and the system was able to identify and recognize the plant assortment and the kind of infection the plant was diseased with up to 100% accuracy.

Indumathi et al. [22] suggested a system for determining the affected leaf region as well as the illness that caused the leaf to be affected. They used K-Medoid clustering and the Random forest algorithm to improve the accuracy of disease detection in leaves. They also used various plant leaf photos. To discover the affected region of the leaf, the image is first pre-processed, and then the clustering technique is used. The system's accuracy is high, as is its ability to predict illnesses. In the system, the time it takes to compute the disease in the infected leaf is minimized, and the memory consumption is therefore manageable.

Bhagat et al. [23] proposed a system for detecting plant leaf disease. SVM is a machine learning algorithm that is being utilized to generate a system for classifying and identifying plant diseases. The method would assist farmers in obtaining a valuable method for precise disease identification with less computational effort. The precision of the SVM classifier was 80%, and when it was combined with Grid Search hyperparameter tuning, it rose to 84%.

Mukhopadhyay et al. [24] for identifying the disease area in tea leaves, image clustering based on the Non-dominated Sorting Genetic Algorithm (NSGA-II) is proposed. Then, for feature reduction and determining the disease in the tea leaves, PCA and multi-class SVM are used, respectively. The results show that the suggested algorithm has an average accuracy of 83% in detecting the type of disease present in tea leaves.

### **3. CLASSIFICATION METHODS**

In this paper, four types of classification methods are used: Decision Tree, Naïve Bayes, Random Forest, and SVM.

## A- DECISION TREE

A basic representation for classifying instances is a Decision Tree. It is a sort of Supervised ML in which data is continually divided depending on a parameter [25][26]. A decision tree is a tree structure with nodes such as the root, intermediate, and leaf nodes. Both classification and regression are done with a decision tree [27][28][29][30].

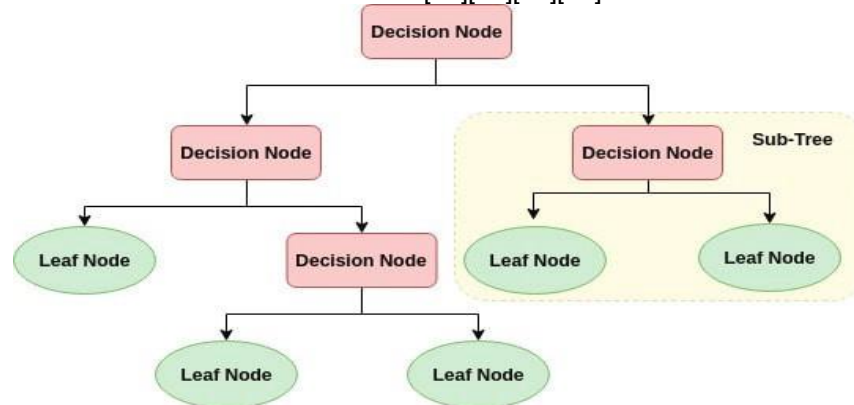


Figure 1 Decision Tree

## B- NAÏVE BAYES

It is a classification algorithm Bayes Theorem-based and the presumption of predictor independence. A Naive Bayes classifier, in concise terms, means that the presence of one attribute in a class is irrelevant to the presence of any other characteristics. For instance, if a fruit has a red color and, round shape, and a diameter of about three inches, it will be called an apple. Even though these features are dependent on one another or the company of other features, they all add to the likelihood that this fruit is an apple, which is the reason why it is called "Naive". The Naive Bayes model is simple to construct and is particularly beneficial for huge data sets. Naive Bayes is recognized to outperform even the most sophisticated classification methods due to its simplicity [31][32][33].

## C- RANDOM FOREST

Random forest is one of the best classifications and regression methods for machine learning. It can be used to classify a huge number of datasets. It consists of a collection of tree-structured classifiers. The tree is based on the sampled random values and the woodland. The data is entered at the top of the tree and then moved down the tree. The data is sampled at random, so there are fewer sets. Random forest trees of a random number are used to find the sample class. The randomizing variable determined how the cuts are discovered regularly [34][35].

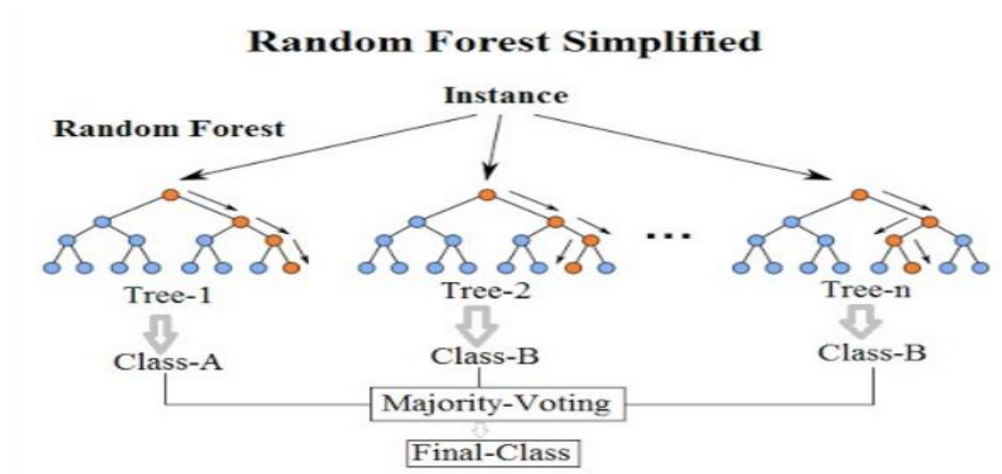


Figure 2 Random Forest

#### D- SUPPORT VECTOR MACHINE

SVM is a supervised ML algorithm that can be utilized for solving issues like classification and regression. It transforms data by using a method identified as the kernel trick, and then discovers an ideal boundary between the possible outputs based on these transformations. Simply put, it performs some extremely difficult data transformations before determining how to distinct data based on the outputs set or labels [36][37].

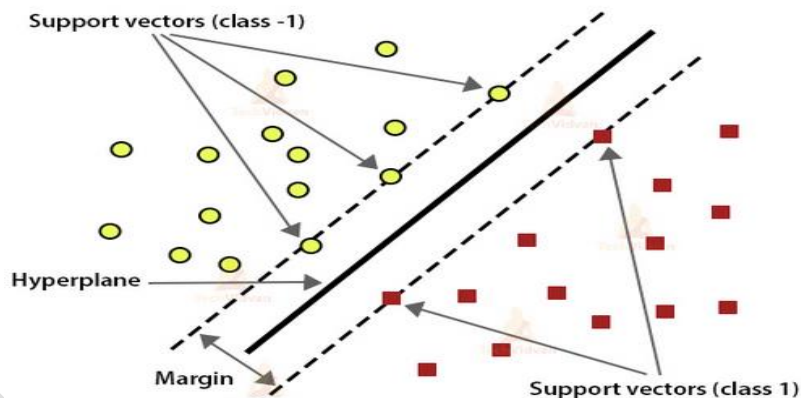


Figure 3 Support Vector Machine

#### 4. PLANT LEAF DISEASES

The diseased plant shows special symptoms and signs that indicate that it has a disease, and the symptoms and signs that appear on the plant may differ depending on the type of disease, the type of host, the degree of resistance to the disease, and the prevailing environmental conditions. Several types of diseases affect plant leaves such as Black Spot, Botrytis Blight, Leaf Spot, Powdery Mildew, and Rust. Most roses may be infected with Black Spot, but some plants are more immune to this disease than others. Many shrubs raised varieties are resistant to this disease, while many hybrid teas raised varieties are extremely susceptible. The black spot occurs on the rose as black splotches that cause leaf yellowing around them [39] Botrytis Blight comes in a variety of ways, the most popular of which flourish in wet, rainy weather. But for the roots, Botrytis Blight will invade all areas of the plant. Brown plant tissue, blight bumps or cankers, and fruit or bulb rot or end rot are all common symptoms. Leaf Spot is a general descriptor for a variety of plant diseases. Leaf

spot is also known as anthracnose or canker in larger clusters or rows [40]. Infected leaves may turn yellow and can drop prematurely. The fungal blight will live and replicate in the plant's dead debris. Powdery mildew spreads on the leaf surface and is usually white or grey. When it first surfaces, it is often mistaken for dust or ashes. Any of the powder will rub off when touched. Powdery mildew outbreaks may have mild to serious consequences depending on the type of plant [41]. Rust is present on plants in over 5000 different animals. They are fungi that fly across the air before landing on a host plant. Defoliation, stunted growth, and branch dieback are some of the most common symptoms of rust [42]. Figure 4 shows plant disease types



**Figure 4 Plant Leaf Diseases**

## 5. METHODS

In this paper Weka tool is used, which is a platform with a graphical user interface that helps the user to directly access the dataset. It is mainly used for classifications and clustering algorithms. Weka is evaluated with five classifiers: tree-based decision tree and random forest, rules-based decision table and PART, and Bayesian-based naïvebayes. In this study a plant leaf dataset also has been used that contains images from various infected and uninfected leaves, Images are gathered from different farms. Information about conditions of each leaf in those images has been stored in an Excel file then the file is saved with CSV file format Since CSV file only contains data about the file names and the condition of each instance, this data must be converted to some numerical format for the prediction algorithms to be applied. We are taken all these data from the Kaggle website and it has been used as input to the existing algorithms.

## 6. RESULTS AND DISCUSSION

After applying those four classification techniques on the plant leaf dataset in the weka tool, the output of the classifier is measured in terms of correctly classified instances, incorrectly classified instances, Kappa statistic, mean absolute error, root mean squared error, relative absolute error, and the root relative squared error. The detailed result of each technique is shown in table 1 and table 2.

**Table 1: The Summary Result for Each Classification Techniques**

	Decision-Tree	Random-Forest	Naïve-Bayes	SVM
<b>Correctly classified</b>	253	232	242	255

<b>instances</b>				
<b>Incorrectly classified instances</b>	57	78	68	55
<b>Kappa Statistics</b>	0.5621	0.2859	0.55	0.5696
<b>Mean absolute error</b>	0.2057	0.355	0.2205	0.1774
<b>Root mean squared error</b>	0.3742	0.3918	0.4352	0.4214
<b>Relative absolute error</b>	47.0084	81.1348	50.4059	40.5537
<b>Root relative squared error</b>	80.0442	83.8155	93.0932	90.1054
		Number of instances	310	

**Table 2: Detailed Result for Each Classification Techniques**

	<b>TP rate</b>	<b>FP rate</b>	<b>Precision</b>	<b>Recall</b>	<b>F-Measure</b>	<b>MCC</b>	<b>ROC Area</b>	<b>PRC Area</b>
<b>Decision-Tree</b>	0.816	0.276	0.812	0.816	0.812	0.566	0.838	0.823
<b>Random-Forest</b>	0.748	0.518	0.794	0.748	0.693	0.389	0.898	0.905
<b>Naïve-Bayes</b>	0.781	0.167	0.826	0.781	0.788	0.575	0.880	0.877
<b>SVM</b>	0.823	0.294	0.820	0.823	0.815	0.577	0.764	0.751

To find the accuracy of those four classification techniques for detecting plant leaf diseases the given formula is used. Accuracy is defined as the number of correctly classified instances divided by the total number of instances multiplied by 100.

$$\text{Accuracy} = \frac{\text{Number of correctly classified instance}}{\text{Total number of instance}} * 100$$

The decision tree classifier has an accuracy of 81.6%, the naive Bayes classifier has an accuracy of 78.1%, the random forest classifier has an accuracy of 74.8%, and the SVM has an accuracy of 82.3%. Based on the result the SVM classifier has the highest accuracy that its correctly classified instance is 255 form 310 total instances, while the random forest has the lowest accuracy. The TP rate and kappa statistics of SVM also is higher than other methods in contrast; the random forest has the lowest TP rate. Based on correctly classified instance the difference between SVM and Decision Tree is only two points/instance that means there is no significant difference between them. Furthermore, SVM has shown the maximum MCC value, unlikely the Random Forest has a minimum MCC value.

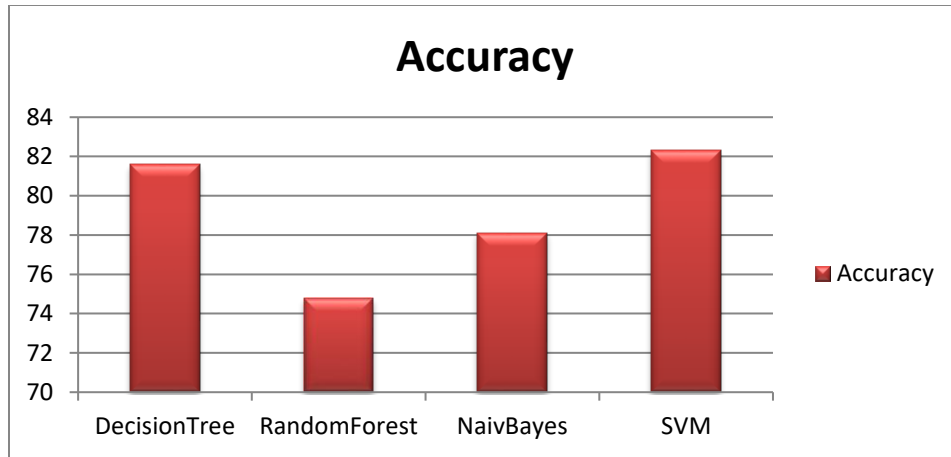


Figure 5 Comparison Accuracy

The execution time taken by the Naïve Bayes algorithm is 0.01 seconds, Decision Tree with 0.15 seconds, SVM algorithm with 0.58 seconds contrary to Random Forest that has taken much more time for execution which is 0.69 seconds as shown in figure 6. Random forest was an unfavorable choice to be used in plant leaf disease detection. Because its accuracy was lower, and its execution time was much more compared to other algorithms which have been used in this paper.



Figure 6 Execution Time

## 7. CONCLUSION

Agriculture plays an important role in our life. It provides us with basic needs such as food and fodder. Also, it develops our economy and power. But unfortunately, many diseases threaten the growth of plants and leaves. Those plants have to be protected and do not let die. For battling these diseases technology can be used to detect and identify them. In this paper a comparison among these classifiers (decision tree, naive Bayes, random forest, and SVM) has been done to detect diseases from plant leaves. The results have illustrated that SVM is the best one due to its high accuracy. As a result, SVM demonstrates that it has the potential to be a highly effective and efficient classifier algorithm.

## REFERENCES

- [1] P. B. Padol and A. A. Yadav, "SVM classifier based grape leaf disease detection," in 2016 Conference on Advances in Signal Processing (CASP), Pune, India, Jun. 2016, pp. 175–179, doi: 10.1109/CASP.2016.7746160.



- [2] D. Q. Zeebaree, H. Haron, A. M. Abdulazeez, and S. R. M. Zeebaree, "Combination of K-means clustering with Genetic Algorithm: A review," vol. 12, no. 24, p. 8, 2017.
- [3] R. G. de Luna, E. P. Dadios, and A. A. Bandala, "Automated Image Capturing System for Deep Learning-based Tomato Plant Leaf Disease Detection and Recognition," in TENCON 2018 - 2018 IEEE Region 10 Conference, Jeju, Korea (South), Oct. 2018, pp. 1414–1419, doi: 10.1109/TENCON.2018.8650088.
- [4] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, "A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction," JASTT, vol. 1, no. 2, pp. 56–70, May 2020, doi: 10.38094/jastt1224.
- [5] T. N. Tete and S. Kamlu, "Detection of Plant Disease Using Threshold, K- Mean Cluster and ANN Algorithm," p. 4.
- [6] S. B. Mohammed and A. M. Abdulazeez, "Deep Convolution Neural Network for Facial Expression Recognition," p. 9, 2021.
- [7] U. B. Korkut, O. B. Gokturk, and O. Yildiz, "Detection of plant diseases by machine learning," in 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, May 2018, pp. 1–4, doi: 10.1109/SIU.2018.8404692.
- [8] M. P. Vaishnave, K. S. Devi, P. Srinivasan, and G. A. P. Jothi, "Detection and Classification of Groundnut Leaf Diseases using KNN classifier," in 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, Mar. 2019, pp. 1–5, doi: 10.1109/ICSCAN.2019.8878733.
- [9] S. S. Kumar and B. K. Raghavendra, "Diseases Detection of Various Plant Leaf Using Image Processing Techniques: A Review," in 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India, Mar. 2019, pp. 313–316, doi: 10.1109/ICACCS.2019.8728325.
- [10] C. G. Dhaware and K. H. Wanjale, "A modern approach for plant leaf disease classification which depends on leaf image processing," in 2017 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, Jan. 2017, pp. 1–4, doi: 10.1109/ICCCI.2017.8117733.
- [11] R. M. Prakash, G. P. Saraswathy, G. Ramalakshmi, K. H. Mangaleswari, and T. Kaviya, "Detection of leaf diseases and classification using digital image processing," in 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, Mar. 2017, pp. 1–4, doi: 10.1109/ICIIECS.2017.8275915.
- [12] M. Islam, AnhDinh, K. Wahid, and P. Bhowmik, "Detection of potato diseases using image segmentation and multiclass support vector machine," in 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE), Windsor, ON, Apr. 2017, pp. 1–4, doi: 10.1109/CCECE.2017.7946594.
- [13] D. QaderZeebaree, A. MohsinAbdulazeez, D. AsaadZebari, H. Haron, and H. NuzlyAbdullHamed, "Multi-Level Fusion in Ultrasound for Cancer Detection based on Uniform LBP Features," Computers, Materials & Continua, vol. 66, no. 3, pp. 3363–3382, 2021, doi: 10.32604/cmc.2021.013314.
- [14] H. Genitha C., E. Dhinesh, and A. Jagan, "Detection of Leaf Disease Using Principal Component Analysis and Linear Support Vector Machine," in 2019 11th International Conference on Advanced Computing (ICoAC), Chennai, India, Dec. 2019, pp. 350–355, doi: 10.1109/ICoAC48765.2019.246866.
- [15] M. B. Ahmad Supian, H. Madzin, and E. Albahari, "Plant Disease Detection and Classification Using Image Processing Techniques: a review," in 2019 2nd International Conference on Applied Engineering (ICAE), Batam, Indonesia, Oct. 2019, pp. 1–4, doi: 10.1109/ICAE47758.2019.9221712.
- [16] E. Hossain, Md. F. Hossain, and M. A. Rahaman, "A Color and Texture Based Approach for the Detection and Classification of Plant Leaf Disease Using KNN Classifier," in 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), Cox'sBazar, Bangladesh, Feb. 2019, pp. 1–6, doi: 10.1109/ECACE.2019.8679247.

- [17] S. Kumar, K. Prasad, A. Srilekha, T. Suman, B. P. Rao, and J. N. Vamshi Krishna, "Leaf Disease Detection and Classification based on Machine Learning," in 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), Bengaluru, India, Oct. 2020, pp. 361–365, doi: 10.1109/ICSTCEE49637.2020.9277379.
- [18] G. Kuricheti and P. Supriya, "Computer Vision Based Turmeric Leaf Disease Detection and Classification: A Step to Smart Agriculture," in 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, Apr. 2019, pp. 545–549, doi: 10.1109/ICOEI.2019.8862706.
- [19] A. H. Bin Abdul Wahab, R. Zahari, and T. H. Lim, "Detecting diseases in Chilli Plants Using K-Means Segmented Support Vector Machine," in 2019 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC), Singapore, Jul. 2019, pp. 57–61, doi: 10.1109/ICISPC.2019.8935722.
- [20] B. Rajesh, M. V. SaiVardhan, and L. Sujihelen, "Leaf Disease Detection and Classification by Decision Tree," in 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), Tirunelveli, India, Jun. 2020, pp. 705–708, doi: 10.1109/ICOEI48184.2020.9142988.
- [21] S. V. Militante, B. D. Gerardo, and N. V. Dionisio, "Plant Leaf Detection and Disease Recognition using Deep Learning," in 2019 IEEE Eurasia Conference on IOT, Communication and Engineering (ECICE), Yunlin, Taiwan, Oct. 2019, pp. 579–582, doi: 10.1109/ECICE47484.2019.8942686.
- [22] R. Indumathi., N. Saagari., V. Thejuswini., and R. Swarnareka., "Leaf Disease Detection and Fertilizer Suggestion," in 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, Mar. 2019, pp. 1–7, doi: 10.1109/ICSCAN.2019.8878781.
- [23] M. Bhagat, D. Kumar, I. Haque, H. S. Munda, and R. Bhagat, "Plant Leaf Disease Classification Using Grid Search Based SVM," in 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, Feb. 2020, pp. 1–6, doi: 10.1109/IDEA49133.2020.9170725.
- [24] S. Mukhopadhyay, M. Paul, R. Pal, and D. De, "Tea leaf disease detection using multi-objective image segmentation," *Multimed Tools Appl*, vol. 80, no. 1, pp. 753–771, Jan. 2021, doi: 10.1007/s11042-020-09567-1.
- [25] O. Ahmed and A. Brifcani, "Gene Expression Classification Based on Deep Learning," in 2019 4th Scientific International Conference Najaf (SICN), Al-Najef, Iraq, Apr. 2019, pp. 145–149, doi: 10.1109/SICN47020.2019.9019357.
- [26] D. Q. Zeebaree, H. Haron, and A. M. Abdulazeez, "Gene Selection and Classification of Microarray Data Using Convolutional Neural Network," in 2018 International Conference on Advanced Science and Engineering (ICOASE), Duhok, Oct. 2018, pp. 145–150, doi: 10.1109/ICOASE.2018.8548836.
- [27] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *JASTT*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.
- [28] A. S. Eesa, Z. Orman, and A. M. A. Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems," *Expert Systems with Applications*, vol. 42, no. 5, pp. 2670–2679, Apr. 2015, doi: 10.1016/j.eswa.2014.11.009.
- [29] A. S. Eesa, A. M. Abdulazeez, and Z. Orman, "A DIDS Based on The Combination of Cuttlefish Algorithm and Decision Tree," *SJUOZ*, vol. 5, no. 4, p. 313, Dec. 2017, doi: 10.25271/2017.5.4.382.
- [30] N. Najat and A. M. Abdulazeez, "Gene clustering with partition around mediods algorithm based on weighted and normalized mahalanobis distance," in 2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Okinawa, Nov. 2017, pp. 140–145, doi: 10.1109/ICIIBMS.2017.8279707.

- [31] F. Harahap, A. Y. N. Harahap, E. Ekadiansyah, R. N. Sari, R. Adawiyah, and C. B. Harahap, "Implementation of Naïve Bayes Classification Method for Predicting Purchase," in 2018 6th International Conference on Cyber and IT Service Management (CITSM), Parapat, Indonesia, Aug. 2018, pp. 1–5, doi: 10.1109/CITSM.2018.8674324.
- [32] Salim, Nareen O. M. and Abdulazeez, Adnan Mohsin, "Human Diseases Detection Based On Machine Learning Algorithms: A Review," Jan. 2021, doi: 10.5281/ZENODO.4467510.
- [33] D. A. Zebari, D. Q. Zeebaree, A. M. Abdulazeez, H. Haron, and H. N. A. Hamed, "Improved Threshold Based and Trainable Fully Automated Segmentation for Breast Cancer Boundary and Pectoral Muscle in Mammogram Images," IEEE Access, vol. 8, pp. 203097–203116, 2020.
- [34] Abdulkareem, Nasiba Mahdi and Abdulazeez, Adnan Mohsin, "Machine Learning Classification Based on Radom Forest Algorithm: A Review," Jan. 2021, doi: 10.5281/ZENODO.4471118.
- [35] M. Govardhan and V. M B, "Diagnosis of Tomato Plant Diseases using Random Forest," in 2019 Global Conference for Advancement in Technology (GCAT), BANGALURU, India, Oct. 2019, pp. 1–5, doi: 10.1109/GCAT47503.2019.8978431.
- [36] U. Mokhtar, M. A. S. Ali, A. E. Hassenian, and H. Hefny, "Tomato leaves diseases detection approach based on Support Vector Machines," in 2015 11th International Computer Engineering Conference (ICENCO), Cairo, Egypt, Dec. 2015, pp. 246–250, doi: 10.1109/ICENCO.2015.7416356.
- [37] D. M. Abdulqader, A. M. Abdulazeez, and D. Q. Zeebaree, "Machine Learning Supervised Algorithms of Gene Selection: A Review," vol. 62, no. 03, p. 13, 2020.
- [38] I. Ibrahim and A. Abdulazeez, "The Role of Machine Learning Algorithms for Diagnosing Diseases," JASTT, vol. 2, no. 01, pp. 10–19, Mar. 2021, doi: 10.38094/jastt20179.
- [39] H.-J. Yu and C.-H. Son, "Leaf Spot Attention Network for Apple Leaf Disease Identification," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, Jun. 2020, pp. 229–237, doi: 10.1109/CVPRW50498.2020.00034.
- [40] Z. Xiuhua, J. Renbao, C. Siyu, J. Yongcheng, and L. Xiaomin, "Study on Pretreatment Technology of Image of Disease Leaf Spot," in 2018 International Conference on Smart Grid and Electrical Automation (ICSGEA), Changsha, Jun. 2018, pp. 173–176, doi: 10.1109/ICSGEA.2018.00051.
- [41] V. Gupta, N. Sengar, M. K. Dutta, C. M. Travieso, and J. B. Alonso, "Automated segmentation of powdery mildew disease from cherry leaves using image processing," in 2017 International Conference and Workshop on Bioinspired Intelligence (IWOBI), Funchal, Jul. 2017, pp. 1–4, doi: 10.1109/IWOBI.2017.8006454.
- [42] A. P. Marcos, N. L. Silva Rodovalho, and A. R. Backes, "Coffee Leaf Rust Detection Using Convolutional Neural Network," in 2019 XV Workshop de Visão Computacional (WVC), São Bernardo do Campo, Brazil, Sep. 2019, pp. 38–42, doi: 10.1109/WVC.2019.8876931.

UNDER PEER REVIEW