Research Article
https://doi.org/10.33904/ejfe.1415402

Eur J Forest Eng 2024, 10(2):84-91

Classification of Powdery Mildew Disease Symptoms on Sandalwood Using Machine Learning Techniques

A. Muthu Kumar[1], J. Cruz Antony[2*], V. Soundararajan[1]

[1]Institute of Wood Science and Technology, Bengaluru, India
[2]Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, Chennai, India

Abstract

Powdery mildew (*Oidium sp*.) is a fungal disease that infects plants by creating white powdery spots on plants and trees, reducing in yield. Powdery mildew is often influenced by changes in climatic conditions with cloud factors, humidity, and temperature playing major roles. This study focuses on building a Machine learning model to classify powdery mildew disease symptoms on sandalwood trees based on abiotic features like soil moisture, temperature, humidity, and cloud factors. Various machine learning algorithms such as Decision Tree, Logistic Regression, Random Forest, Support Vector Machine, and K-Nearest Neighbors were used on the dataset, and the model with the highest accuracy was chosen for building a powdery mildew prediction web application on the cloud platform. This web application helps in the prediction of the disease incidence/intensity and thereby enlightens the farming community to adopt appropriate management strategies.

Keywords: IWST, Machine learning, Powdery mildew, Sandalwood.

1. Introduction

Sandalwood (*Santalum album*), commonly known as Indian sandalwood, is a small tropical tree native to the dry regions of Southern India. From its native area of distribution, the *S. album* was introduced to Northern, Central (Srinivasan et al., 1992), Northwestern (Das and Jagatpati, 2013), and North-Eastern India (Viswanath, 2014; Babita et al., 2018). The tremendous extension in sandalwood cultivation was only in the past decade; before that, sandalwood was largely confined to the forests and government-owned plantations of Karnataka, Kerala, and Tamil Nadu. In the early 2000s, the Karnataka and Tamil Nadu governments brought a policy change in the liberalization of growing sandalwood, which allowed many private owners and firms to cultivate sandalwood (Pallavi and Patel, 2015). Historically, sandalwood is blended with Indian culture and heritage (Fox, 2000). Sandalwood oil is the main product extracted from the heartwood of the sandalwood tree to produce perfumes, candles, incense sticks, soaps, and religious and cultural purposes (Subasinghe, 2013). It is growing and reviving naturally under suitable conditions in India as it is a part of indigenous vegetation (Rai, 1990).

However, several factors, such as diseases and pests, now threaten the growth of sandalwood. More than 150 pest species appear on sandalwood trees, but only a few induce severe damage, which includes defoliators, stem borers, sapsuckers, and termites. Sandalwood planted at nurseries is greatly affected by seedling disease and is considered a severe threat causing economic loss (Das, 2021). Plant and tree species are frequently exposed to pathogens like bacteria, viruses, and fungi, leading to severe crop yield loss globally (Anderson et al., 2004). Studies on diseases of sandalwood in recent times are lacking; however, there was a report on canker symptoms of sandalwood in Karnataka (Nagaveni et al., 2014), powdery mildew symptoms in Madhya Pradesh and Maharashtra (Patel et al., 2015; Chirame, 2018; Bankar et al., 2019), and blight (Muthu Kumar et al., 2021a). Powdery mildews are a crucial group of plant pathogenic fungi composing a white, powdery film on stems, leaves, flowers, and fruits of Angiosperms (Braun and Cook, 2012). In most cases, the early symptoms of powdery mildew are hard to spot, as signs of infection appear small on budding leaves and raised blisters formed out of the conidia (asexual spore). The symptoms in sandalwood are highly favored by prevailing environmental/abiotic conditions (Muthu Kumar et al., 2021b). In general, the symptoms appeared as discrete, cobweb-like to powdery white patches of hyphae (Figure 1).

As the symptomatic condition progresses, the white spot-like structure will grow larger, irregular shape, and are largely interconnected. The findings in the study are interconnected with symptom expression in *Acacia mangium*, as the infection progressed, these white patches increased in size and coalesced to form larger

*Corresponding Author: Tel: +91 8147274846 E-mail: jcruzantony@gmail.com
Received: 19 January 2024; Accepted: 31 May 2024

Figure 1. Symptoms of powdery mildew incidence

patches, spreading to newly developed phyllodes (Borah et al., 2012). The symptom expression is highly feasible under conducive environmental conditions (Muthu Kumar et al., 2021b). The powdery nature of the disease helps in the fast dissemination of the conidia; thereby, the impact of the loss is very severe. Appropriate diagnosis of the symptom and data analysis facilitate the prevention and control of the disease.

Data analysis comprising traditional statistical methods is focused on finding relationships between factors, whereas machine learning techniques perform numerous tasks, including classification and prediction (Rajula et al., 2020). Decision Tree and Random Forest, the popular machine learning models, were used to predict tomato leaf diseases such as mosaic virus, yellow curl, septoria spot, and bacterial spot. Multiple features were fused with the ML model to improve computational time and accuracy. The proposed decision tree method yielded an accuracy of 90%, and the random forest model gave 94% classification accuracy, but the model is limited to fewer disease classifications (Basavaiah and Anthony, 2020). Popular supervised ML techniques such as Decision Tree, Naïve Bayes, Support Vector Machine, K-Nearest Neighbor, and Random Forest were used to detect maize plant diseases. Out of five models, the Random Forest model has performed well with a classification accuracy of 79.23%, and the model is limited to classifying only the maize disease dataset (Panigrahi et al. 2020). The k-nearest neighbor model was used to classify rice leaf diseases, viz., Brown spot, Bacterial leaf blight, and Leaf smut. The k-NN model used the GLCM feature extraction method, achieving an accuracy of 68.83% with a 0.485 kappa value was achieved (Saputra et al. 2020). To classify groundnut leaf diseases such as *Cercospora arachidicola*, *Phaeoisariopsis personatum*, *Puccinia arachidis*, and Bud Necrosis virus, which causes heavy destruction of groundnut crops, the k-NN classification model was built upon the dataset. The k-NN model performed well when compared with the existing SVM model (Vaishnnave et al., 2019).

The demand for early disease warnings and a surge in biological data leads to the necessity of opting for machine learning techniques. These techniques are capable of generating significant contributions to classifying various biological problems (Martinelli et al., 2015). Early prediction of powdery mildew is crucial in attaining good quality yield, and Machine Learning algorithms come in handy for the classification of the disease in accordance with the weather conditions. Machine learning models are considered economical compared to the employing of chemical pesticides. Various ML classifiers like Naïve Bayes, Support Vector Machine, and K Nearest Neighbor were used to predict 15 plant diseases using the WEKA tool. With a classification accuracy of 95%, the KNN classifier outperformed the other ML classifiers (Prathusha et al., 2019). Several ML classifiers were used in disease classification, and the Gaussian support vector machine showed an accuracy of 94.74%, which was better compared to other ML classifiers (Bhatia et al., 2021). A smart decision support system incorporating these ML classifiers is required for early-stage plant disease diagnosis to regulate and prevent losses (Sannakki et al., 2011). These types of support systems, especially web-based ones, are an asset to the non-experts in the identification and classification of plant diseases (Pertot et al., 2012).

Hence, early detection and treatment of powdery mildew is crucial due to its rapid development (Bhatia et al., 2020), which affects sandalwood growth. The hypothesis of the present study involves building various popular ML classification models, which include Decision Tree, Logistic Regression, Random Forest, Support Vector Machine, and K-Nearest Neighbors using the dataset obtained from the factors influencing the symptoms associated with powdery mildew and the best-performing model with good classification accuracy was chosen to build an application and host it on the cloud platform. Finally, the overall outcome of the hypothesis is the web-based ML application that assists the stakeholders in predicting symptoms associated with powdery mildew.

2. Materials and Methods

Data analysis plays a vital role in the interpretation of hypotheses, wherein statistical tools are commonly used for analysis purposes. Machine learning techniques have an advantage over statistical methods by focusing on the data themselves and highlighting the performance of the given task. Various frequently used machine learning techniques used in research are Artificial Neural Networks (ANN), Support Vector Machine (SVM), Naïve Bayesian, and Random Forest (Tahmooresi et al., 2018; Ghorbanzadeh et al., 2019).

Building a successful machine learning model involves preprocessing, analyzing, and fitting the best classifier model and evaluation. Each step in this process is crucial as it contributes to the accuracy of classifying powdery mildew diseases based on the patterns identified by the classifier model during the ML process. The architecture diagram of the proposed system to build the web application of the powdery mildew classification model is depicted in Figure 2.

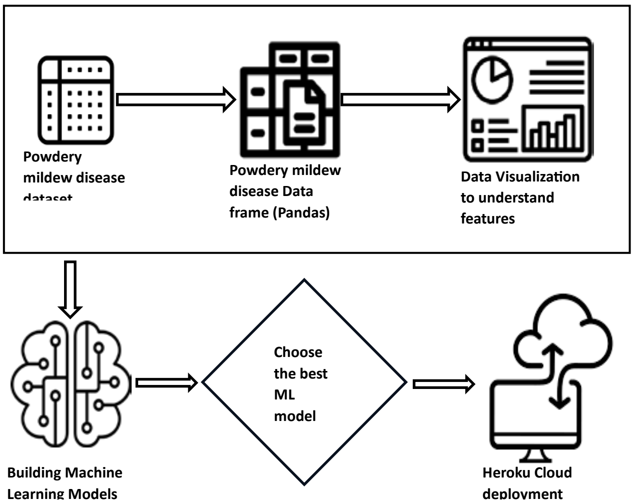

Figure 2. Architecture of the proposed system

2.1. Dataset

The dataset consists of powdery mildew symptoms (class/target feature) on sandalwood from IWST Nursery classified as High, Medium, and Low, along with its respective abiotic features viz., Cloud factor (Smith et al., 2017), Soil moisture (Jaing et al., 2021), Temperature (Singh et al., 2023), and Humidity (Romero et al., 2021) from the year 2016 to 2018. The abiotic features, along with the target feature, were pre-processed by handling the missing and outlier data values, and the abiotic features were discretized (sample dataset mentioned in Table 1) as per the economic threshold level (Viswanath and Chakraborty, 2022).

Table 1. Discretized abiotic features according to the Economic Threshold Level

Abiotic features	Discrete values	Actual value (range)
Cloud factor	Broken	3 to 7 Oktas
	Closed	8 to 9 Oktas
	Open	0 to 2 Oktas
Soil moisture	Adequate	70 to 90 %
	Inadequate	below 70% (or 50 to 70%)
Temperature	Maximum	30^{0}C to 35^{0}C
	Minimum	15^{0}C to 25^{0}C
	Moderate	25^{0}C to 30^{0}C
Humidity	Maximum	90% to 100%
	Minimum	70% to 90%
	Moderate	50% to 70%

2.2. Powdery mildew disease data analysis

Data analysis helps analyze the data structure, find hidden patterns, visualize the effects of one feature over the others, and evaluate the data. Pandas is a Python package that provides fast, flexible, and expressive data structures. It aims to be the fundamental high-level building block for practical, real-world data analysis in Python (McKinney, 2015). The dataset proposed in this study was transferred into the Pandas data frame for data analysis. The Python libraries such as Seaborn and Matplotlib were used to analyze the powdery mildew disease dataset with its abiotic features. Matplotlib is a Python package used for visualizing data through graphs and charts, making it a key component in data science. Seaborn extends Matplotlib for creating beautiful graphics using a more straightforward set of methods. It is also more integrated with Pandas data frames (Bisong, 2019).

The count plot between cloud factors with respect to the target feature and symptoms (Figure 3) depicts that powdery mildew incidences are high when the cloud is broken and closed, similarly, the symptoms are medium with the open cloud. The count plot between soil moisture with respect to the symptoms portrays that powdery mildew incidences are high when soil moisture is adequate. Comparatively, data is limited for adequate soil moisture (Figure 4). The count plot of temperature with respect to the symptoms (Figure 5) depicts the occurrence of powdery mildew as high when the temperature is moderate. When the temperature is modest the symptoms are moderate. The count plot between humidity with respect to the symptoms illustrates that the symptoms are high when the humidity is maximum and moderate; similarly, the symptoms are medium when the humidity is minimum (Figure 6).

Symptoms Cloud factor	High	Low	Medium
Broken	222	1	44
Closed	125	3	103
Open	101	12	277

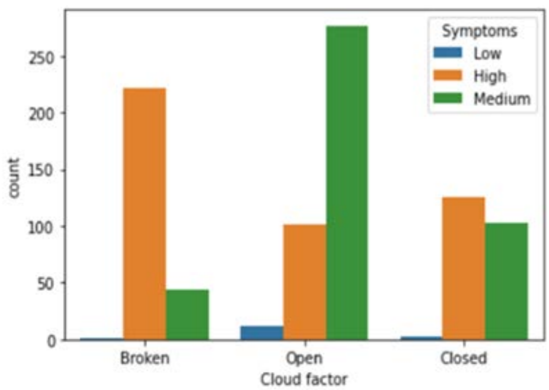

Figure 3. Count plot of Cloud factor with Symptoms

Symptoms	High	Low	Medium
Soil moisture			
Adequate	448	16	420
Inadequate	0	0	4

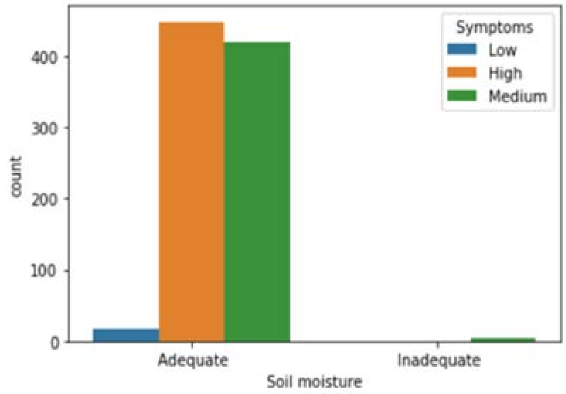

Figure 4. Count plot of Soil moisture with Symptoms

Symptoms	High	Low	Medium
Temperature			
Maximum	42	0	43
Minimum	99	2	243
Moderate	307	14	138

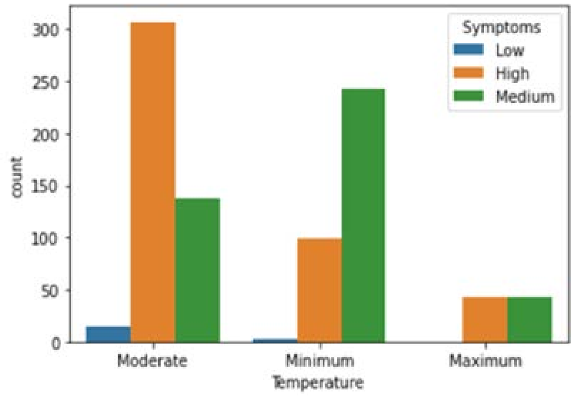

Figure 5. Count plot of Temperature with Symptoms

Symptoms	High	Low	Medium
Humidity			
Maximum	183	16	9
Minimum	20	0	262
Moderate	245	0	153

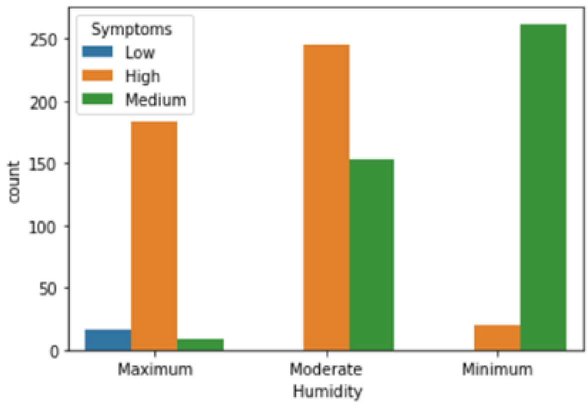

Figure 6. Count plot of Humidity with Symptoms

A correlation heatmap was generated using Python's Matplotlib library (Figure 7), is a graphical representation of the correlation matrix revealing the correlation between features in the powdery mildew dataset. The correlation matrix depicted in Figure 7 depicts that the cloud factor has a strong correlation with the target feature, powdery mildew symptoms. Climatic seasonality and climate change were relevant agents of plant disease epidemics, where some minor diseases have become primary diseases in the subcontinent (Sharma et al. 2007).

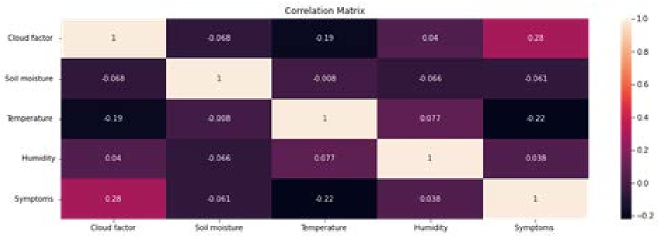

Figure 7. Correlation matrix of powdery mildew symptoms and abiotic features.

2.3. Powdery mildew disease ML models

Machine Learning (ML) models have significantly emerged to build new opportunities to resolve and understand the thorough process of data used in agricultural environments (Liakos et al., 2018). The ML models have the ability to identify patterns from the historical dataset and try to classify the new problem without human intervention. Usually, the ML classification model pursues two stages; in the first stage, the classification model tries to fit, i.e., get trained based on the past data where the target feature is known, and in the next stage, the model tries to classify the target feature with the aid of the trained model. Hence, the dataset was split into training and test datasets with 80% and 20% respectively (Banjare et al., 2021).

Five frequently used machine learning algorithms for the classification of biological data were involved, and their performance was evaluated based on their classification accuracy (Mousavizadegan and Mohabatkar, 2016). The scikit-learn machine learning library was imported, and the ML algorithms viz., Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), and K Nearest Neighbor (KNN) were used to build the powdery mildew disease ML classification models. The LR (shortly called as logit model) is frequently used for predictive modelling and analysis. The LR model recognizes the relationship between the dependent and independent variables by determining logistic regression probability (Ranganathan et al., 2017). The SVM is considered a powerful model among ML models, identifying exquisite patterns from a complex dataset. The SVM model builds a decision boundary, i.e., a hyperplane between two classes that aids the classification of labels from feature vectors (Huang et al., 2018). DTs are multivariate and can effectively handle complex datasets with a facile parametric design. DT model classifies a population into an inverted tree-like structure and has become popular in multidisciplinary research (Song and Ying, 2015). The RF is a successful classification algorithm that merges different random decision trees and aggregates by calculating the mean of the classifications. RFs have displayed good performance, especially when the number of attributes is huge compared to the number of tuples (Biau and Scornet, 2016). KNN is extensively used in classification problems. KNN works based on the Euclidean distance function that measures the correlation between two instances (Jiang et al., 2007).

3. Results and Discussion

Studying the powdery mildew disease dynamics on sandalwood was successfully achieved by training various machine learning models. The next important step is to validate these trained models with the aid of test data and measure their accuracy in the classification of the disease. The accuracy of the ML classifiers is usually estimated with the aid of a confusion matrix. This matrix provides insight into how well the classifier has identified the multiclass population. The True Negative (TN), True Positive (TP), False Negative (FN), and False Positive (FP) of the actual and predicted population instances were calculated from which the confusion matrices were derived for the five ML classification models (Figure 8).

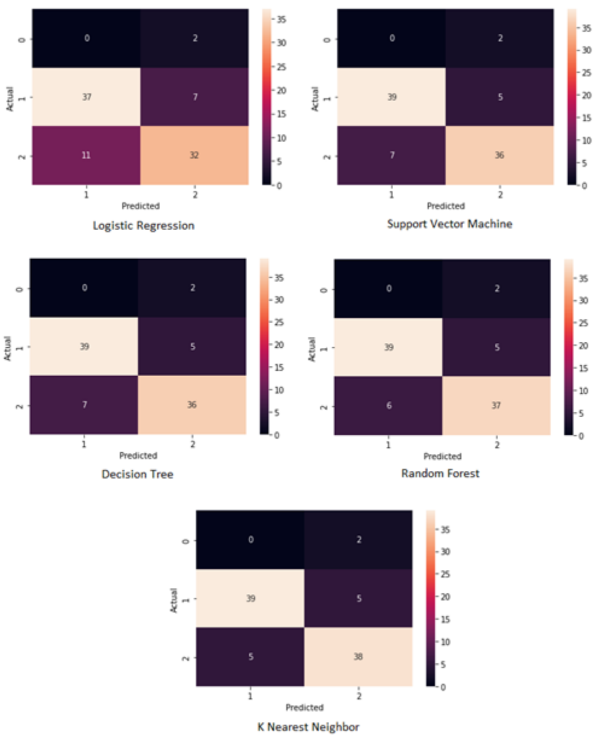

Figure 8. Confusion matrices of ML classification models.

Further, the classification accuracy for each classifier was obtained by calculating the ratio of the total diagonal values and total values in the confusion matrix (TP+TN)/(TP+TN+FN+FP). The classification accuracy of the trained ML models in predicting the instances of the powdery mildew dataset is given in Figure 9.

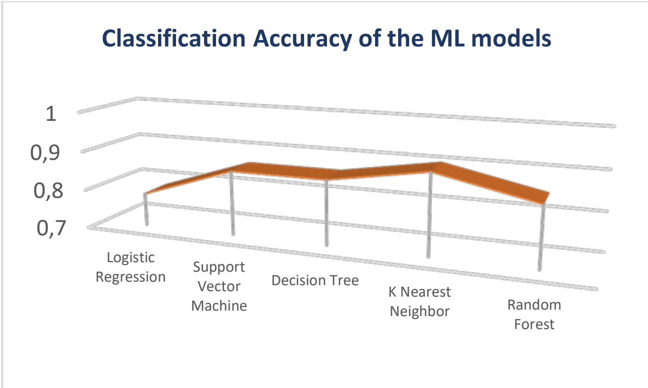

Figure 9. Classification accuracy of the ML models.

The graph (Figure 9) clearly shows that the K Nearest Neighbor (KNN) classifier has displayed better accuracy (90.23%) than the other ML models. Based on that, the KNN machine learning web application for predicting the powdery mildew symptoms was developed using the Streamlit Python library.

Technology and the rise of advanced techno approaches like machine learning have been vastly implemented to distinguish diseases in plants and are considered non-destructive disease recognition methods (Abdu et al., 2020). The application of trained machine learning models has been proven successful in studying the powdery mildew disease dynamics on sandalwood.

The proposed trained model using test data was validated, and the accuracy of the ML classifiers was estimated with the aid of a confusion matrix (Antony and Pratheepa, 2017; Ashwin et al., 2021). It was based on the actual and predicted population instances confusion matrices derived for the five ML classification models. The classification accuracy (Patro and Patra, 2014) of the trained ML models in predicting the instances of the powdery mildew dataset revealed that the K Nearest Neighbor (KNN) classifier displayed better accuracy (90.23%) when compared with the other ML models. The KNN is elemental and intuitive and it classifies unknown data instances based on identical and neighbor-classified instances (Mucherino et al., 2009). KNN is easily adaptable to new training samples and thus increases the overall accuracy rate of the classifier model (Liao and Vemuri, 2002).

Nowadays, mobile and server-based techniques have become popular in disease prediction (Ramesh et al., 2018) Hence the KNN-based Machine learning web application for predicting powdery mildew symptoms was developed using the Streamlit Python library. Streamlit is an easy-to-use Python library designed to develop interactive web applications from basic Python code with the help of Streamlit Cloud (Parker et al., 2021).

4. Conclusion

Powdery mildew disease is one of the predominant diseases in sandalwood nurseries. In addition to the proximity of plants, abiotic factors like cloud cover, soil moisture, temperature, and humidity instigate the symptoms associated with the pathogen *Pseudoidium santalacearum*. The conducive environment favours the pathogen and spreads swiftly from one plant to the other. Understanding the prevailing abiotic factors helps in the diagnosis of the symptoms and prediction of the disease. Based on the abiotic features from the dataset, the classification of powdery mildew symptoms on sandalwood using five important machine learning classifier models was performed. The K Nearest Neighbor classifier has shown a better classification accuracy of 90.23% among all the classifier models. The ML model accuracy can still be increased by adding more training data, as the KNN can conform to new training data records. Further, the classification model that used the KNN classifier was chosen to develop a powdery mildew symptoms prediction web application (https://github.com/ummefahad/Powdery-ldewclassification.git). Web application is considered an interactive way to disseminate knowledge about the powdery mildew dynamics on sandalwood ro the public/researcher. The Streamlit Python library, a handy tool, was used to build interactive web application. It uses Streamlit Cloud to deploy the classification model Python code. The application also provides expert information on the precautionary steps as well as the treatment of powdery mildew diseases on sandalwood.

Ethics Committee Approval: N/A.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept: M.K.A., J.C.A.; Design: J.C.A., M.K.A., S.V.; Supervision: M.K.A., J.C.A.; Resources: M.K.A., S.V.; Data Collection: M.K.A., S.V.; Analysis: J.C.A., M.K.A.; Literature Search: J.C.A., M.K.A.; Writing Manuscript: M.K.A., J.C.A., S.V.; Critical Review: M.K.A., J.C.A., S.V.

Conflict of Interest: The authors have no conflicts of interest to declare.

Financial Disclosure: The authors declared that this study has received no financial support.

Cite this paper as: Kumar, A.M., Antony, J.C., Soundararajan, V. 2024. Classification of Powdery Mildew Disease Symptoms on Sandalwood Using Machine Learning Techniques, *European Journal of Forest Engineering*, 10(2):84-91.

References

Abdu, A.M., Mokji, M.M., Sheikh, U.U. 2020. Machine learning for plant disease detection: an investigative comparison between support vector machine and deep learning. *IAES International Journal of Artificial Intelligence*, 9:670-683.

Anderson, P.K., Cunningham, A.A., Patel, N.G., Morales, F.J., Epstein, P.R., Daszak, P. 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. Trends in ecology & evolution, 19(10):535-544.

Annabel, L.S.P., Annapoorani, T. and Deepalakshmi, P. 2019. Machine Learning for Plant Leaf Disease Detection and Classification–A Review. In *2019 International Conference on Communication and Signal Processing (ICCSP)* (pp. 0538-0542). IEEE.

Antony, J.C. and Pratheepa, M. 2017. Study of population dynamics of soybean semi-looper *Gesonia gemma* Swinhoe by using rule induction model in Maharashtra, India. *Legume Research-An International Journal*, 40(2):369-373.

Ashwin, N., Adusumilli, U.K., Kemparaju, N., Kurra, L. 2021. A machine learning approach to prediction of soybean disease. *International Journal of Scientific Research in Science, Engineering and Technology*, 9:78-88.

Babita, M., Sandeep, C., Sushant, A., Sruthi, S., Syam, V. 2018. Assessment of heartwood and oil content of Santalum album Linn. in natural and naturalized populations across contrasting edapho-climatic conditions in India. *Indian Forester*, 144(7):675-685.

Banjare, P., Matore, B., Singh, J., Roy, P.P. 2021. In silico local QSAR modeling of bioconcentration factor of organophosphate pesticides. *In Silico Pharmacology*, 9(1):1-13.

Bankar, P., Kadam, V., Bhosale, A., Shitole, S., Wagh, S., Chandankar, S., Chitale, R. and Kanade, M.B. 2019. Powdery mildew fungi from Phaltan Area of Satara District, Maharashtra. *Int J Curr Microbiol App Sci*, *8*(7):2181-6.

Basavaiah, J., and Arlene Anthony, A. 2020. Tomato leaf disease classification using multiple feature extraction techniques. Wireless Personal Communications, 115(1):633-651.

Bhatia, A., Chug, A. and Singh, A.P. 2020. February. Hybrid SVM-LR classifier for powdery mildew disease prediction in tomato plant. In *2020 7th International conference on signal processing and integrated networks (SPIN)* (pp. 218-223). IEEE.

Bhatia, A., Chug, A., Singh, A.P. 2021. Statistical analysis of machine learning techniques for predicting powdery mildew disease in tomato plants. *International Journal of Intelligent Engineering Informatics*, *9*(1):24-58.

Biau, G. and Scornet, E. 2016. A random forest guided tour. *Test*, *25*(2):197-227.

Bisong, E. 2019. Matplotlib and Seaborn. In: *Building Machine Learning and Deep Learning Models on Google Cloud Platform*. Apress, Berkeley, CA.

Borah, R.K., Gogoi, J., Gogoi, B., Sharma, G.S. 2012. New record of Powdery Mildew on *Acacia mangium* Willd. In India. *Journal of Plant Protection Research*, *52*(1):64-66.

Braun, U. and Cook, R.T.A. 2012. Taxonomic manual of Erysiphales (powdery mildews). *CBS Biodiversity Series*, 11.

Chirame, B.B. 2018. First report of powdery mildew of *Ageratum houstonianum*, Mentha and Sandalwood in Maharashtra (India). *Electronic International Interdisciplinary Research Journal*, 7(12):12-14.

Christopher, J.S., Jamie, M.B., Rolf, C. 2017. Cloud cover effect of clear-sky index distributions and differences between human and automatic cloud observations, *Solar Energy*, 144:10-21, ISSN 0038-092X, https://doi.org/10.1016/j.solener.2016.12.055.

Das, S.C. and Jagatpati, T. 2013. Effect of GA3 on seed germination of sandal (Santalum album L.). *International Journal of Current Science*, 8:79-84.

Das, S.C. 2021, Diseases and Insect Pests of Sandalwood. In: Pullaiah, T., Das, S.C., Bapat, V.A., Swamy, M.K., Reddy, V.D., Murthy, K.S.R. (eds) Sandalwood: Silviculture, Conservation and Applications. Springer, Singapore. https://doi.org/10.1007/978-981-16-0780-6_9

Fox, J.E. 2000. Sandalwood: the royal tree. *Biologist (London, England)*, 47(1):31-34.

Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S.R., Tiede, D., Aryal, J. 2019. Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. *Remote Sensing*, 11(2):196.

Huang, S., Cai, N., Pacheco, P.P., Narrandes, S., Wang, Y., Xu, W. 2018. Applications of support vector machine (SVM) learning in cancer genomics. *Cancer genomics & proteomics*, *15*(1):41-51.

Jiang, L., Cai, Z., Wang, D., Jiang, S. 2007. August. Survey of improving k-nearest-neighbor for classification. In *Fourth international conference on fuzzy systems and knowledge discovery (FSKD 2007)* (Vol. 1, pp. 679-683). IEEE.

Jiang, G., Wang, N., Zhang, Y. et al., 2021. The relative importance of soil moisture in predicting bacterial wilt disease occurrence. *Soil Ecology Letters*, 3:356–366. https://doi.org/10.1007/s42832-021-0086-2

Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D. 2018. Machine learning in agriculture: A review. *Sensors*, 18(8):2674.

Liao, Y. and Vemuri, V.R. 2002. Use of k-nearest neighbor classifier for intrusion detection. *Computers & Security*, *21*(5):439-448.

Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., Villa, P., Stroppiana, D., Boschetti, M., Goulart, L.R., Davis, C.E. 2015. Advanced methods of plant disease detection. A review. *Agronomy for Sustainable Development*, 35(1):1-25.

Mckinney, W. 2015. Pandas—Powerful Python Data Analysis Toolkit, 1625.

Mousavizadegan, M. and Mohabatkar, H. 2016. An evaluation on different machine learning algorithms for classification and prediction of antifungal peptides. *Medicinal Chemistry*, 12(8):795-800.

Mucherino, A., Papajorgji, P.J., Pardalos, P.M. 2009. K-nearest neighbor classification. In *Data mining in agriculture* (pp. 83-106). Springer, New York, NY.

Muthu Kumar, A., Soundararajan, V., Rekha J., Divya Bharathi, M., Mamatha N. 2021a. Foliar blight disease of Indian Sandalwood (Santalum album l.) trees caused by the pathogen *Pestalotiopsis guepinii* (desm.) Steyaert. *Asian Journal of Microbiology, Biotechnology & Environmental Sciences*, 23(3):96-100.

Muthu Kumar, A., Soundararajan, V., Manoj Kumar P.E., Abhilash A., Rekha J., Vijayalakshmi, G., Mamatha N. 2021b. Impact of change in microclimatic factors for incidence and prevalence of powdery mildew disease in Indian sandalwood (*Santalum album* L.). *Ecology, Environment and Conservation*, 28(1):271-273.

Nagaveni, H.C., Sundararaj, R., Vijayalakshmi, G. 2014. First report of canker disease on Indian sandalwood (Santalum album Linn.) in India. *Journal on New Biological Reports*, 3(2):120-124.

Pallavi, A. and Patel, A. 2015. Return of scented wood. *Down to Earth*. https://www.downtoearth.org.in/coverage/forests/return-of-scented-wood-48569

Panigrahi, K.P., Das, H., Sahoo, A.K., Moharana, S. C. 2020. Maize leaf disease detection and classification using machine learning algorithms. In Progress in Computing, Analytics and Networking: Proceedings of ICCAN 2019 (pp. 659-669). Springer Singapore.

Parker, A., Heflin, A., Jones, L.C. 2021. Analyzing University of Virginia Health publications using open data, Python, and Streamlit. *Journal of the Medical Library Association: JMLA*, *109*(4):688.

Patel, R.P., Pandey, G.N., Tiwari, G., Patidar, H., Patidar, D.K. 2015. First report of Powdery mildew fungi on Sandal wood in Madhya Pradesh. *Int J Curr Res*, *7*(6):16705-16708.

Patro, V.M. and Patra, M.R. 2014. Augmenting weighted average with confusion matrix to enhance classification accuracy. *Transactions on Machine Learning and Artificial Intelligence*, 2(4):77-91.

Pertot, I., Kuflik, T., Gordon, I., Freeman, S., Elad, Y. 2012. Identificator: A web-based tool for visual plant disease identification, a proof of concept with a case study on strawberry. *Computers and electronics in agriculture*, 84:144-154.

Prathusha, P., Murthy, K.S., Srinivas, K. 2019. Plant Disease Detection Using Machine Learning Algorithms. In *International Conference on Computational and Bio-Engineering* (pp. 213-220). Springer, Cham.

Rai, S.N. 1990. Status and cultivation of sandalwood in India. *In: Hamilton, Lawrence; Conrad, C. Eugene, technical coordinators. Proceedings of the Symposium on Sandalwood in the Pacific;* April 9-11, 1990; Honolulu, Hawaii. Gen. Tech. Rep. PSW-GTR-122. Berkeley, CA: Pacific Southwest Research Station, Forest Service, US Department of Agriculture, 122: 66-71.

Rajula, H. S. R., Verlato, G., Manchia, M., Antonucci, N., Fanos, V. 2020. Comparison of conventional statistical methods with machine learning in medicine: diagnosis, drug development, and treatment. Medicina, 56(9):455.

Ramesh, S., Hebbar, R., Niveditha, M., Pooja, R., Shashank, N., Vinod, P.V. 2018 April. Plant disease detection using machine learning. In *2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C)* (pp. 41-45). IEEE.

Ranganathan, P., Pramesh, C.S. and Aggarwal, R. 2017. Common pitfalls in statistical analysis: logistic regression. *Perspectives in clinical research*, 8(3):148.

Romero, F., Cazzato, S., Walder, F., Vogelgsang, S., Bender, S.F., Van Der Heijden, M.G.A. 2022. Humidity and high temperature are important for predicting fungal disease outbreaks worldwide. *New Phytologist*, 234:1553–1556.

Sannakki, S.S., Rajpurohit, V.S., Nargund, V.B., Kumar, A.R. and Yallur, P.S. 2011. A hybrid intelligent system for automated pomegranate disease detection and grading. *Int. J. Mach. Intell*, 3(2):36-44.

Saputra, R.A., Wasiyanti, S., Saefudin, D.F., Supriyatna, A., Wibowo, A. 2020. Rice leaf disease image classifications using KNN based on GLCM feature extraction. In Journal of Physics: Conference Series (Vol. 1641, No. 1, p. 012080). IOP Publishing.

Sharma, R.C., Duveiller, E., Ortiz-Ferrara, G. 2007. Progress and challenge towards reducing wheat spot blotch threat in the Eastern Gangetic Plains of South Asia: is climate change already taking its toll? *Field Crops Research*, 103(2):109-118.

Singh, B.K., Delgado-Baquerizo, M., Egidi, E. et al., 2023, Climate change impacts on plant pathogens, food security and paths forward. *Nature Reviews Microbiology*, 21:640–656. https://doi.org/10.1038/s41579-023-00900-7.

Song, Y.Y. and Ying, L.U. 2015. Decision tree methods: applications for classification and prediction. *Shanghai archives of psychiatry*, 27(2):130.

Srinivasan, V.V., Sivaramakrishnana, V.R., Rangaswamy, C.R., Anathapadmanabha, H.S., Shankaranarayana, K.H. 1992. Sandal- (Santalum album Linn.). Published by Institute of Wood Science and Technology, Indian Council of Forestry Research and Education, Dehra Dun, India, p.233.

Subasinghe, S.M.C.U.P. 2013, Sandalwood research: a global perspective. *Journal of Tropical Forestry and Environment*, 3(1):1-8.

Tahmooresi, M., Afshar, A., Rad, B.B., Nowshath, K.B., Bamiah, M.A. 2018. Early detection of breast cancer using machine learning techniques. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 10(3-2):21-27.

Vaishnnave, M.P., Devi, K.S., Srinivasan, P., Jothi, G.A.P. 2019. Detection and classification of groundnut leaf diseases using KNN classifier. In 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN) (pp. 1-5). IEEE.

Viswanath, S. 2014. Sandalwood, An unexplored treasure. *Inside Fact, National monthly*, 1(2):10-13.

Viswanath, S., and Chakraborty., 2022. Indian Sandalwood Cultivation Prospects in India. In: Arunkumar, A.N., Joshi, G., Warrier, R.R., Karaba, N.N. (Eds.) Indian Sandalwood. Materials Horizons: From Nature to Nanomaterials. Springer, Singapore.

Yang, X. and Guo, T. 2017. Machine learning in plant disease research. *European Journal of BioMedical Research*, 3(1):6-9.