IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

MACHINE LEARNING BASED EARLY DETECTION OF YELLOW LEAF DISEASE IN ARECANUT PLANT USING SOIL SAMPLES

[1]Ms. Shwetha Kamath, [2]Prakhyath Devadiga, [3]Prathiksha G K, [4]Prathiksha S, [5]Shrinidhi Shervegar

[1]Assitant Professor, Mangalore Institute of Technology and Engineering, Moodabidre, Karnataka, India
[2] Student, Mangalore Institute of Technology and Engineering, Moodabidre, Karnataka, India
[3]Student, Mangalore Institute of Technology and Engineering, Moodabidre, Karnataka, India
[4]Student, Mangalore Institute of Technology and Engineering, Moodabidre, Karnataka, India
[5]Student, Mangalore Institute of Technology and Engineering, Moodabidre, Karnataka, India

Abstract: Arecanut is an important cash crop grown extensively in India, especially in the southern states. It is also known as betel nut and is a major source of income for many farmers in the region. Arecanut is consumed widely in the form of chewing and is an integral part of social and religious practices in India. However, one of the major threats to arecanut cultivation is the Yellow Leaf Disease (YLD). The disease is characterized by yellowing of leaves, stunted growth, and reduced yield, leading to significant economic losses for farmers. The disease is also highly contagious and can spread rapidly, making early detection crucial for effective mitigation measures. Traditionally, the detection of YLD has relied on visual symptoms, which can be challenging to identify in the early stages. This delay in detection can result in significant crop losses. Therefore, there is a need for an early detection system that can identify the disease at an early stage and help farmers take necessary measures to control its spread. In this context, machine learning-based solutions using soil samples have emerged as a promising approach for early detection of YLD in arecanut plants. By analyzing various soil parameters, ID3 algorithm can detect the presence of the disease at an early stage. This can help farmers take timely action to prevent the spread of the disease and improve crop yields.

Keywords - Yellow Leaf Disease, Arecanut, Machine Learning, Soil parameters, ID3 Algorithm.

INTRODUCTION

Arecanut is a significant cash crop that is extensively cultivated in several parts of the world, especially in India, Indonesia, and the Philippines. However, Arecanut plantations are threatened by several fungal diseases, including the Yellow Leaf Disease (YLD). YLD is a devastating fungal disease that affects the Arecanut plant's growth and productivity, leading to significant economic losses for the farmers.

The early detection of YLD is crucial to control the disease's spread and prevent economic losses. Traditional methods of YLD detection rely on visual inspection, which is time-consuming and not always reliable. Therefore, there is a need for a more reliable and cost-effective method for the early detection of YLD.

In this project, we propose a machine learning-based approach for the early detection of YLD in Arecanut plants using soil samples. The proposed method will use soil samples collected from the Arecanut plantations to identify the presence of the fungal pathogen responsible for YLD. Machine learning algorithms and statistical techniques will be used to analyze the soil samples and predict the probability of YLD infection in the Arecanut plants.

The project's main objective is to provide a reliable and cost-effective early detection system for YLD in Arecanut plants, which can be used by farmers to prevent the spread of the disease and improve their yields. The proposed system will be based on the analysis of soil samples, which can provide early indications of the presence of YLD before any visible symptoms appear in the Arecanut plants. Overall, the project aims to contribute to the development of sustainable and efficient agricultural practices by utilizing the power of machine learning and data analysis.

NEED OF THE STUDY

Yellow Leaf Disease (YLD) is a common and devastating fungal disease affecting the Arecanut plant, which is a major cash crop in many parts of the world. The disease causes yellowing of leaves, stunting of growth, and a decrease in the yield of the plant. Early detection of the disease is crucial to prevent the spread of the disease and minimize the economic losses incurred by the farmers.

The project aims to develop a machine learning-based model for the early detection of Yellow Leaf Disease in Arecanut plants using soil samples. The proposed model will use soil samples collected from the affected plants to identify the presence of the fungal pathogen responsible for the disease. The model will use various machine learning algorithms and statistical techniques to analyze the soil samples and predict the probability of YLD infection in the Arecanut plants.

The project will involve the collection of soil samples from Arecanut plantations and the extraction of relevant features from the soil samples. These features will include various physical and chemical parameters of the soil, such as pH, moisture content, and nutrient levels, as well as the presence of specific fungal pathogens. The extracted features will be used to train the machine learning model, which will then be used to predict the likelihood of YLD infection in the Arecanut plants based on the soil samples.

The ultimate goal of the project is to provide a reliable and cost-effective early detection system for YLD in Arecanut plants that can be used by farmers to prevent the spread of the disease and improve their yields.

LITERATURE REVIEW

1.Kanan, L. V et al., (2021). Arecanut Yield Disease Forecast using IoT and Machine Learning. International Journal of Scientific Research in Engineering & Technology, 2(2), 11-15.

This paper presents a system to predict if an arecanut plant is susceptible to disease infection by comparing data values from IoT sensors and historical research. A range is determined for scoring results and a value greater than 8 indicates a high possibility of disease infection. The output is manually confirmed with experience, helping to evaluate the model efficiently. The farmers are warned about the possibility of disease and can take necessary precautions. The model can be extended to multiple crops and diseases based on thorough research.

2.Anilkumar M G., Karibasaveshwara TG., Pavan HK., Sainath Urankar., Dr. Abhay Deshpande. (2021). Detection of Diseases in Arecanut Using Convolutional Neural Networks. International Research Journal of Engineering and Technology (IRJET), 8(5), 4282-4286.

They used Convolutional Neural Network (CNN) a Deep Learning used to diagnose arecanut illnesses. It takes an image as input, assigns learnable weights and biases to distinct items in the image, and then learns to distinguish one from the other based on the results. To train and evaluate the CNN model, they built a dataset of 620 photos of healthy and ill arecanuts. The train and test data are split in an 80:20 split. For model compilation, categoricalcross-entropy is utilized as the loss function, with Adam as the optimizer function and accuracy as the metrics. A total of 50 Epochs are used to train the model to achieve high validation and test accuracy with minimum loss.

3.Huang, K. Y. (2012). Detection and classification of areca nuts with machine vision. Computers & Mathematics with Applications, 64(5), 739-746.

In this initially the grade of arecanuts was found and determined. To categorize the arecanut diseases, a detection line (DL) approach was employed. Six geometrical features, three-color features, and the fault area were used in the categorization technique.

4. L.Sathish Kumar., Mrs.A.Padmapriya.M.C. ID3 Algorithm Performance of Diagnosis for Common Disease.

This paper discusses the problem of constraining and summarizing two algorithms of data mining used in medical prediction. It focuses on using Decision Tree and Bayesian classification for intelligent and effective common disease diagnosis. The outcome of diagnosis data mining technique on the same dataset reveals that Decision Tree outperforms and Bayesian classification is having similar accuracy as of ID3. This paper proposes a procedure for retrieval of dataset with relevant fields using ID3 algorithm, based on individual diagnosis for specific symptoms of the disease.

5. Mallikarjuna, S B, Palaiahnakote Shivakumara, Vijeta Khare, Vinay Kumar N, Basavanna M, Umapada Pal, and Poornima B. (2021). CNN based method for multi-type diseased arecanut image classification. Malaysian Journal of Computer Science, 34(3), 255-265.

This work proposes a new method for classification of arecanut images of different diseases. It exploits multi-gradient concept to generate four gradient images, which enhance fine details irrespective of disease effect. The proposed method outperforms existing methods in terms of classification rate, recall, precision and F-measure. The future target is to extend the idea for more classes and identify disease by analyzing the content of the images.

6.Jiawei Guo., Yu Jin., Huichun Ye., Wenjiang Huang., Jinling Zhao., Bei Cui., Fucheng Liu., Jiajian Deng. Recognition of Areca Leaf Yellow Disease Based on Planet Scope Satellite Imagery

This article compares field crops like wheat to areca yellow leaf disease remote sensing surveillance. The RF model, which outperformed the BPNN and AdaBoost models by 2.95% and 20.59%, had the greatest overall recognition accuracy for areca yellow leaf disease, according to the results (88.24%). The geographic distribution map's improved recognition accuracy made it more useful for areca yellow leaf disease detection and monitoring.

METHODOLOGY

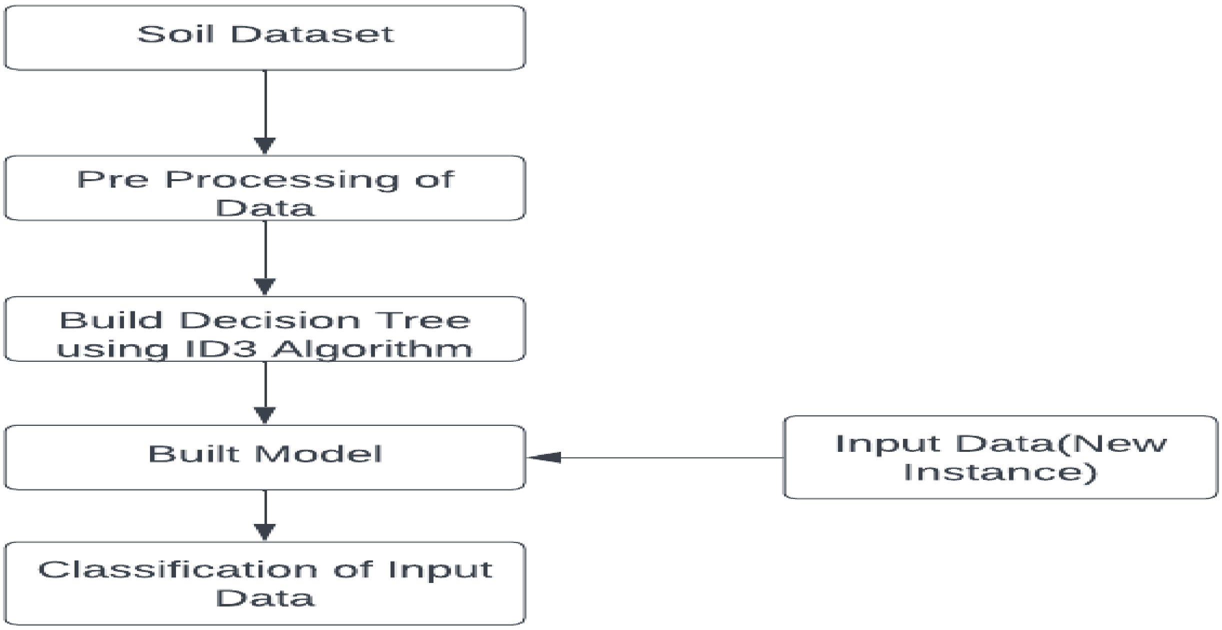

Fig 1. Methodology of early prediction of yellow leaf disease

The soil dataset would typically consist of various attributes or features related to soil characteristics. These features can be used as input variables to train a machine learning model to predict the occurrence of YLD in Arecanut plants.

Soil dataset features that may be relevant for detecting YLD:

- pH level: Soil pH is a measure of its acidity or alkalinity, and it can influence the availability of nutrients in the soil. Some pathogens, including the one that causes YLD, may prefer certain pH levels for optimal growth. Therefore, pH level could be an important feature in the dataset.

- Micronutrients content: Micronutrients are essential elements required by plants in small quantities for their growth and development. Imbalanced or deficient micronutrient levels in the soil can affect plant health and make them more susceptible to diseases, including YLD.

- Nutrient levels: Soil nutrient levels, such as nitrogen, phosphorus, and potassium, play a critical role in plant growth and development. Imbalanced nutrient levels can weaken plant resistance to diseases, including YLD.

- Organic matter content: The organic matter content of the soil is an important indicator of soil health. Soil with higher organic matter content is generally more fertile and supports healthier plant growth. Organic matter content can influence the availability of nutrients and microbial activity in the soil, which can impact the occurrence of YLD.

The specific features and their importance may vary depending on the context, location, and factors affecting the occurrence of YLD in the target area.

Preprocessing of data involves converting numerical values to the discrete values to feed to csv file. Discrete feature values play a crucial role in feature selection, tree construction, tree pruning, and classification in the ID3 algorithm, which is a widely used decision tree learning algorithm for machine learning tasks.

The ID3 (Iterative Dichotomiser 3) algorithm is a popular decision tree algorithm used for building decision trees for classification tasks. Here's how it works:

- Selecting the Root Node: The ID3 algorithm starts by selecting the best feature to use as the root node of the decision tree. The feature selection is based on the concept of information gain, which measures the reduction in entropy (a measure of impurity) of the dataset after splitting on a particular feature. The feature with the highest information gain is selected as the root node.

- Splitting the Dataset: Once the root node is selected, the dataset is split into subsets based on the values of the selected feature. Each subset contains samples with the same value for the selected feature.

- Repeating the Process: The ID3 algorithm recursively repeats the above steps for each subset created in the previous step. The algorithm continues to split the dataset based on the remaining features with the highest information gain until one of the following conditions is met:

1.All samples in the subset have the same class label, resulting in a leaf node with that label.
2.There are no more features left to split on, resulting in a leaf node with the majority class label of the remaining samples.

- Pruning: Once the decision tree is fully constructed, it may be pruned to prevent overfitting. This can be done using various techniques such as pre-pruning (limiting the depth of the tree) or post-pruning (pruning unnecessary branches after tree construction).

- Prediction: To make predictions with the decision tree, new samples are recursively traversed down the tree based on the feature values, and the majority class label of the samples in the leaf node reached is used as the predicted class label for the input sample.

Once the decision tree is built the input instances can be provided to model which is being built to predict the presence of yellow leaf disease in arecanut.

RESULTS AND DISCUSSION

Fig 2. Home Page

The input page contains the input instances that needs to be considered to predict the yellow leaf disease. The attributes include Micronutrients along with pH, N, P, K values. Once the user input is given the result page provides with the result of the prediction. The result may be “Yes” or “No” based on the ID3 algorithm which builds decision tree and the input provided.

Fig 3. Input Page

Fig 4. Result Page

Fig 5. Result Page

CONCLUSION

The project highlights the potential in providing a solution for the early detection of Yellow Leaf Disease in Arecanut plants. The project involves several steps, including the collection of soil samples from various sources, categorizing the samples based on several factors such as pH, moisture, and nutrients, and preprocessing the data.

The ID3 algorithm is then used to train a model to predict the occurrence of Yellow Leaf Disease in Arecanut plants using soil samples. The algorithm's advantage is that it can handle both categorical and continuous data and is a powerful tool for decision-making problems. The model's performance is evaluated using a test dataset, and the parameters are optimized to achieve the best possible accuracy. ID3 algorithm has a fast training and prediction time, making it suitable for large datasets and real-time applications such as detection of yellow leaf disease in arecanut plant. It uses a top-down approach to decision-making, which is based on selecting the best attribute to split the data.

The project's potential impact is significant, as early detection and prevention of Yellow Leaf Disease can result in better crop yields and improved economic outcomes for farmers. The project can be deployed as an early warning system for the disease, allowing farmers to take action to mitigate its effects.

In summary, the project has the potential to reduce the impact of Yellow Leaf Disease on Arecanut plants, improve productivity, and provide economic benefits to farmers.

REFERENCES

[1] Kanan, L. V et al., (2021). Arecanut Yield Disease Forecast using IoT and Machine Learning. International Journal of Scientific Research in Engineering & Technology, 2(2), 11-15

[2] Anilkumar M G., Karibasaveshwara TG., Pavan HK., Sainath Urankar., Dr. AbhayDeshpande. (2021). Detection of Diseases in Arecanut Using Convolutional Neural Networks.

[3] Huang, K. Y. (2012). Detection and classification of areca nuts diseases with machine vision.Computers & Mathematics with Applications, 64(5), 739-746.

[4] L.Sathish Kumar, Mrs.A.Padmapriya.M.C ID3 Algorithm Performance of Diagnosis for Common Disease.

[5] Mallikarjuna, S B, Palaiahnakote Shivakumara, Vijeta Khare, Vinay Kumar N, Basavanna M, Umapada Pal, and Poornima B. (2021). CNN based method for multi-type diseased arecanut image classification. Malaysian Journal of Computer Science, 34(3), 255-265.

[6] Jiawei Guo., Yu Jin., Huichun Ye., Wenjiang Huang., Jinling Zhao., Bei Cui., Fucheng Liu., Jiajian Deng. Recognition of Areca Leaf Yellow Disease Based on Planet Scope Satellite Imagery