International Journal of Engineering Applied Sciences and Technology, 2024
Vol. 9, Issue 01, ISSN No. 2455-2143, Pages 147-151
Published Online May 2024 in IJEAST (http://www.ijeast.com)

VALIDATION OF ARECA NUT FRUIT ROT DISEASE DATASET USING WEATHER PARAMETERS

Dr. Jothimani K
Professor, Department of Master of Computer Applications
Vivekananda College of Engineering and Technology
Puttur, India

P M Venugopala
II MCA,
Department of Master of Computer Applications
Vivekananda College of Engineering and Technology
Puttur, India

Abstract—**Forecasting susceptibility to fruit rot disease in areca nut through the analysis of weather data using machine learning techniques. The research uses data from the OpenWeatherMap API, with a specific focus on important factors like precipitation, highest temperature, and moisture levels. Research uses machine learning models to predict disease scores and categorize disease severity levels. The information is divided into separate training and testing sets, then utilized to teach the models how to predict disease outcomes using weather data.**

Index Terms—**Areca nut, Fruit rot disease, Decision Tree, Weather**

I. INTRODUCTION

The areca nut is a significant crop in many Southeast Asian countries, especially in India, contributing significantly to the global agricultural sector. India produces more than half of the world's areca nuts, showing its significant economic and cultural significance. However, the susceptibility of areca nut farms to diseases like fruit rot continues to pose a persistent threat to the farmers' livelihoods. Various environmental elements like temperature, precipitation, humidity, and wind can influence the progression and effects of these illnesses. Recognizing the significance of disease control, this research is initiating groundbreaking work to develop a unified weather based prediction model for areca nut diseases using machine learning algorithms.

Integrating machine learning into traditional farming methods has greatly transformed digital agriculture through the inclusion of advanced technologies. Despite the possibility, there has been a shortage of studies on using ML to predict diseases in areca nuts, particularly regarding the weather elements affecting fruit rot. This study aims to fill this gap by introducing a new dataset that integrates historical weather data with cases of illnesses.

The emphasis on predicting diseases based on weather for areca nut fruit rot disease aligns with the pressing need to mitigate the effects of climate variability on crop health. Conventional methods often use image-based models, which restrict their applicability to situations after the disease has already occurred. Alternatively, a predictive model based on weather has the capability to notify farmers of potential disease outbreaks, enabling them to take timely preventive actions.

A. Objective

Research seeks to assist areca nut growers by utilizing machine learning to forecast fruit rot diseases impacted by weather factors such as temperature, precipitation, and humidity. The research will create a novel dataset by merging past weather data with disease cases, allowing for the development and assessment of forecasting models. This extensive dataset will enhance agricultural predictive modeling and boost the precision and dependability of disease forecasting, filling the current void of public domain datasets.

B. Scope

The presence of fruit rot disease, depicted in "Fig. 1, presents a significant danger to the financial well-being of farmers, resulting in significant losses in crops and economic insecurity. This study aims to create a predictive model for fruit rot disease in areca nut plantations using machine learning and weather data. The study plans to improve predictive accuracy by merging disease occurrences with historical weather data to develop a new dataset.

Fig. 1. Disease affected areca nuts

II. RELATED WORKS

Multiple studies have investigated how weather factors are linked to forecasting fruit rot diseases in areca nut. Rajashree Krishna and K. V. Prema [1] studied past weather records in two areas and used a rule-based algorithm to combine weather and disease data, discovering that vanilla GRU and bidirectional LSTM models were most effective in predicting fruit rot diseases. Another research conducted by K. Rajashree, K.V. Prema, G. Rajath, and S. Angad [2] employed a comparable method, merging illness data with historical weather information and evaluating different predictive models, ultimately showcasing the efficiency of diverse techniques.
Additional study called "Creating and Validating Areca Nut Disease Dataset using Machine Learning Techniques based on Weather Parameters" [3] merged weather information from the Zone Agricultural and Horticultural Research Station in Brahmavar, Udupi, with disease data obtained from farmer surveys. The dataset was validated through regression models such as SVR, DTR, RFR, and MLP, with Random Forest Regression (RFR) showing the highest accuracy and the lowest error rate for predicting fruit rot disease.
Regarding disease control, the research paper "Managing Fruit Rot Disease of Arecanut Caused by Phytophthora meadii" [4] suggests the use of Bordeaux mixture as a leaf spray, in addition to cultural methods such as clearing dropped nuts and trimming branches for improved air circulation, as successful approaches to combat fruit rot disease.
Research on crops like cotton and rice, as well as areca nut, has shown the important influence of weather conditions on disease occurrence. A study conducted by Aman Sharma, Sudhir Kumar Mishra, and Harish Kumar [5] showcased how weather impacts bacterial blight of cotton, and another research on "Infection Risk of Bacterial Grain Rot of Rice" [6] emphasized the importance of weather-disease datasets for predicting plant diseases.
Furthermore, a research study dedicated to statistical feature selection [7] highlighted the importance of key factors in impacting variables related to crops, which have an impact on both areca nut yield and tree lifespan. These studies emphasize the significance of combining weather data with disease information to enhance predictive accuracy and formulate efficient management strategies for areca nut and other crops.

III. METHODOLOGY

Ensuring the precision of forecasts for areca nut fruit rot disease severity depends on careful data collection and thorough training procedures. This study thoroughly gathers extensive data sets that include important variables for predicting diseases. These datasets are carefully selected and organized to include key variables necessary for predicting diseases. Using sophisticated statistical techniques and machine learning algorithms, the dataset is methodically taught to recognize patterns that signal the severity of a disease. This methodical process ensures strong outcomes, as shown in “Fig. 2”, demonstrating the step-by-step flow from gathering data to creating predictive models, resulting in accurate predictions of disease presence and intensity in areca nut crops.

A. Dataset Collection

The first step is to make a dataset that combines weather conditions with disease information, which is crucial as there is no publicly available data for areca nut fruit rot disease. Goal of this study is to verify the dataset using weather information from Dakshina Kannada district. Key factors like temperature,

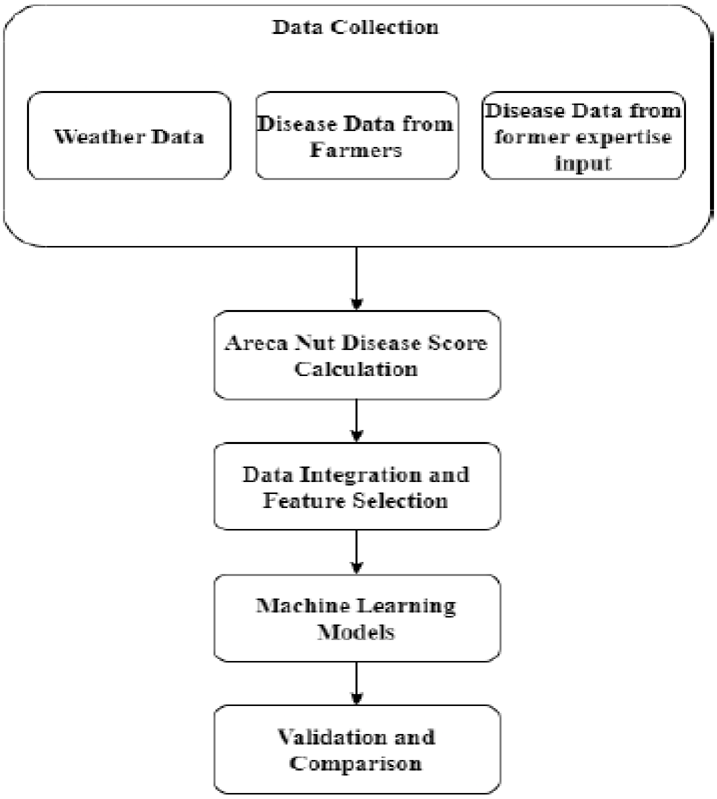

Fig. 2. Disease severity prediction system

humidity, and rainfall play a major role in affecting disease development. Following a comprehensive analysis, an appropriate dataset was found, encompassing data from January 1, 2015, to December 31, 2023, comprising 3,288 data entries stored in a CSV file and including all essential variables. Excess columns were then eliminated to simplify the weather dataset for additional analysis. A sample of dataset is shown in "Fig. 3".

datetime	tempmax	tempmin	temp	feelslikem	feelslikemi	feelslike	dew	humidity
01-01-2015	31.2	23	25.7	33.7	23	26.3	21.1	76.4
02-01-2015	31.1	21.4	25.2	32.1	21.4	25.6	20	73.9
03-01-2015	30.7	22.6	25.7	32.2	22.6	26.1	19.9	71.2
04-01-2015	31.9	21.2	25.5	34	21.2	25.9	19.7	71.4
05-01-2015	31.1	21.3	25.9	32.5	21.3	26.6	20.5	72.9
06-01-2015	30.7	22.9	26.5	33.8	22.9	27.4	21.3	74.3
07-01-2015	31.2	23.3	26.8	35.1	23.3	27.7	21.7	75
08-01-2015	31.1	22.6	26.2	31.8	22.6	26.7	20.1	70.7
09-01-2015	31.1	21.6	26.4	32.3	21.6	26.9	19.7	68.3
10-01-2015	33.1	22.1	26.5	32.6	22.1	27	20	69.1
11-01-2015	32.4	20.1	25.6	30.4	20.1	25.1	13.6	51.4
12-01-2015	32.4	19.7	25.1	30.9	19.7	25	15.3	56.5
13-01-2015	32.1	19.2	24.7	30.8	19.2	24.7	15.8	59
14-01-2015	30.7	19.2	24.8	31.4	19.2	24.9	17	63.2
15-01-2015	31.4	18.9	25.1	31.5	18.9	25.2	17.5	64.3
16-01-2015	32.4	18.8	25.7	31.9	18.8	25.9	18.1	64.5
17-01-2015	32.9	20.1	26.1	33.6	20.1	26.5	19	66.3
18-01-2015	32.1	21	26.6	35.3	21	27.4	20.4	69.9

Fig. 3. Sample of weather dataset

B. Areca Nut Disease Score Calculation

A custom function was created to calculate disease scores for each entry, utilizing a rule-based algorithm and criteria provided by experienced betel nut farmers. These standards, designed for the Dakshina Kannada District, evaluate aspects such as temperature, humidity, and rainfall in order

International Journal of Engineering Applied Sciences and Technology, 2024
Vol. 9, Issue 01, ISSN No. 2455-2143, Pages 147-151
Published Online May 2024 in IJEAST (http://www.ijeast.com)

to measure the seriousness of diseases. Scores rise in situations conducive to illness, like temperatures under 26°C, humidity over 80%, and rainfall over 15mm. On the other hand, ratings go down in unfavorable conditions, such as temperatures exceeding 28°C with humidity less than 70% and rainfall less than 15mm, or temperatures above 35°C with humidity below 50%. The scoring system, with a maximum of 35 points, offers a consistent way to assess disease risk, making it easier to classify and analyze how environmental factors impact disease outbreaks in the area.

C. Machine Learning Models

Afterwards, a distinct CSV file was generated as the training dataset, including weather information (temperature, humidity, precipitation) along with disease ratings and severity levels. Python's sklearn package was used to create three different machine learning models: Random Forest Regression, Linear Regression, and Decision Trees. The primary objective of the project was to create a thorough dataset for confirming fruit rot diseases through weather factors.

D. Validation and Comparision

Currently, the model can forecast the seriousness of fruit rot disease for the upcoming seven days. The user can choose the machine learning model for predictions with the frontend component. Decision Tree Regression was the model with the lowest Mean Squared Error (MSE) and Mean Absolute Error (MAE) among the three models. After choosing a model and clicking the predict button, the model is trained and predicts the disease score and severity class for each row of weather data.

E. Predicting severity of the Disease

Following an assessment of various weather data APIs, OpenWeatherMap was chosen for its extensive features. On clicking "Predict" button in web interface, it gathers weather data depending on latitude and longitude inputs. After the system determines the user's whereabouts, it collects weather information for the next seven days beginning from today. An selected machine learning model subsequently forecasts disease scores for every data point in the dataset. The severity category forecasted is based on the average score over the seven days of weather data.

IV. RESULT AND DISCUSSION

Using the web interface, the user has the option to choose from various models in order to estimate the severity of areca nut fruit rot disease in a specific area. If Random Forest Regression, a type of ensemble learning, is selected for its capability to manage non-linear relationships and high dimensional data.

Linear Regression as a reference point for comparison, despite its simplicity and better interpretability compared to more sophisticated models. It allows for a basic comparison with more complicated methods by establishing a straightforward linear connection between weather conditions and the severity of diseases.

Decision rules based on weather parameters are used in Decision Tree Regression to model disease severity, allowing for interpretation of results. The model's ability to understand complex patterns and relationships between variables allows it to accurately assess how weather conditions affect disease severity. In "Fig. 4", Decision Trees display the smallest MAE and MSE in comparison to the other two models.

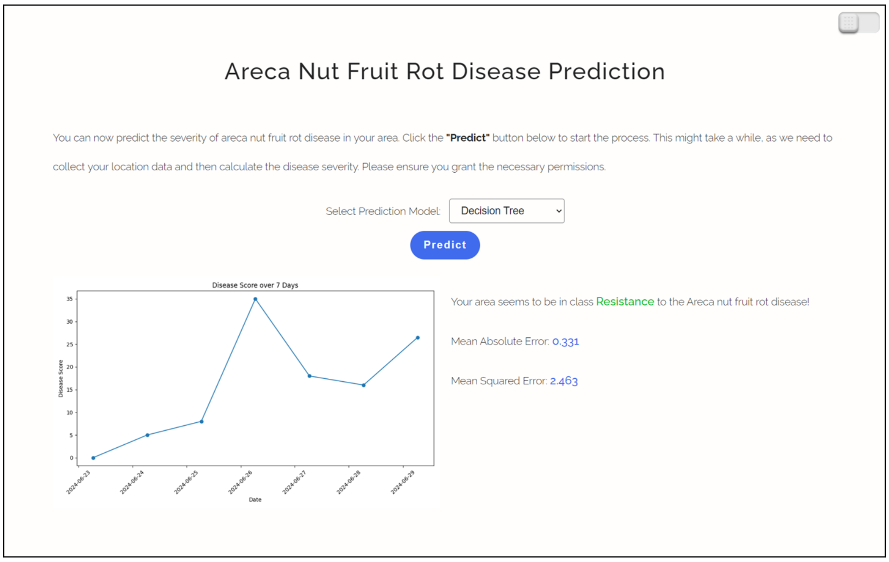

Fig. 4. Output of Decision Tree Regression

Overall, the assessment of the system's effectiveness shows that the Decision Tree Regression and Random Forest Regression models surpass the Linear Regression model in predicting disease severity. The Decision Tree model is preferred for predicting areca nut disease severity using weather variables because of its high accuracy and lower error rates. This thorough examination validates the efficiency and dependability of the suggested system in aiding areca nut farmers in early disease detection and management.

V. CONCLUSION

Research has a significant influence on the agricultural sector, particularly for farmers cultivating areca nuts and encountering issues related to fruit rot disease. The study helps farmers predict disease outbreaks by connecting weather patterns to disease severity, enabling them to take preventive actions based on weather forecasts. Utilizing this proactive strategy may lead to higher agricultural productivity, decreased financial damages, and improved food security. Moreover, the thorough dataset created in this study addresses an essential information gap for the public and acts as a beneficial tool for future research, enhancing the precision and dependability of disease forecasts using machine learning. Furthermore, it was found that incorporating more weather variables into the prediction models can lead to significantly better results. Including factors such as duration of sunlight, speed of wind, and moisture of soil can improve the precision of predictions on disease severity. This recognition highlights the importance of continuously enhancing the model by integrating a broader range of data sources, ensuring more reliable and useful predictions for farmers. These lessons emphasize how the research can enhance strategies for managing diseases, ultimately benefiting the farming community in Dakshina Kannada.

VI. REFERENCES

[1]. Krishna, R., & Prema, K. V. (2023). Constructing and Optimising RNN models to predict fruit rot disease incidence in areca nut crop based on weather parameters. IEEE Access.

[2]. Rajashree, K., Prema, K. V., Rajath, G., & Angad, S. (2022). Prediction of fruit rot disease incidence in Arecanut based on weather parameters.

[3]. Krishna, R., Prema, K. V., & Gaonkar, R. (2022). Areca nut disease dataset creation and validation using machine learning techniques based on weather parameters. Engineered Science, 19, 205-214.

[4]. Narayanaswamy, H., Raju, J., & Jayalakshmi, K. (2017). Management of fruit rot disease of arecanut incited by Phytophthora meadii. Inter. J. Cur. Microbiol. Appl. Sci, 6, 2824-2828.

[5]. SHARMA, A., MISHRA, S. K., & KUMAR, H. (2017). Influence of weather variables on incidence and progress of bacterial blight of cotton in south western Punjab. Editorial Board, 2017, 234.

[6]. Kim, H. S., Do, K. S., Park, J. H., Kang, W. S., Lee, Y. H., & Park, E. W. (2020). Application of numerical weather prediction data to estimate infection risk of bacterial grain rot of rice in Korea. The plant pathology journal, 36(1), 54.

[7]. Pakkala, P. G. R., Rai, B. S., & Rai, P. (2023). Statistical Driven Feature Selection for Prognostic Reasoning and Insight Exploration of Areca Nut Crop using Data Analytics Approach. Indian Journal of Science and Technology, 16(20), 1495-1504.

[8]. SHARMA, A., MISHRA, S. K., & KUMAR, H. (2017). Influence of weather variables on incidence and progress of bacterial blight of cotton in south western Punjab. Editorial Board, 2017, 234.

[9]. Narayanaswamy, H., Raju, J., & Jayalakshmi, K. (2017). Management of fruit rot disease of arecanut incited by Phytophthora meadii. Inter. J. Cur. Microbiol. Appl. Sci, 6, 2824-2828.

[10]. Sunil, K. M., Devadas, V. S., & George, P. S. (2011). Influence of weather parameters on yield and yield attributes of areca nut (Areca catechu L.). Journal of Agricultural Physics, 11, 88-90.

[11]. Balanagouda, P., Sridhara, S., Shil, S., Hegde, V., Naik, M. K., Narayanaswamy, H., & Balasundram, S. K. (2021). Assessment of the spatial distribution and risk associated with fruit rot disease in Areca catechu L. Journal of Fungi, 7(10), 797.

[12]. Pandian, R. T. P., Thube, S. H., Bhavishya, Merinbabu, Chaithra, Santhoshkumar, P., ... & Hegde, V. (2021). First report of Phytophthora palmivora (EJ Butler) EJ Butler, 1919 causing fruit rot in Areca triandra Roxb. ex Buch.-Ham. from India. Australasian Plant Pathology, 50(4), 495-499.