Machine Learning Classification Models for Early Plant Stress Detection

Olayiwola Joy Oluwabukola* & Adamu Dahiru
Department of Computer Engineering, Federal Polytechnic Ilaro, Ogun State, Nigeria
Corresponding author email: oluwabukola.olayiwola@federalpolyilaro.edu.ng,

Introduction

Machine learning approaches have been widely applied in recent times for agricultural research because of their capacity to handle large datasets, recognise subtle patterns, and diagnose stress with minimal human interaction. Early diagnosis of plant stress is important for preserving optimum crop health and production potential since it allows for early management and lowers the risk of disease, insect, or environmental challenges. Traditional plant stress monitoring methods, such as visual inspection and biochemical assays, are labour-intensive, time-consuming, and susceptible to human error. The importance of KNN, decision tree, and random forest classifiers in the detection of plant stress in its early stages is assessed in this study. It examines the effectiveness of using fundamental criteria such as accuracy, precision, recall, and F1-score in determining plant stress. These measures are necessary because they show the model's performance in the detection of the stressed plants with fewer errors, i.e., false positives and false negatives. This work contributes to the expanding body of knowledge on the application of algorithms for learning in agricultural precision farming. The research also provides useful information about selecting appropriate classification models for monitoring plant stress.

Materials and methods

Machine Learning algorithms and environmental sensor technologies introduced an innovative system for plant health prediction. The Plant health classification uses supervised learning techniques with Random Forest (RF), KNN, and decision tree classifier as the main methods. The predictive technique includes four steps that lead to selecting the most efficient prediction model after training and validating it. The system aims to enhance diagnosis accuracy through usage of the curated dataset for live monitoring combined with predictive analytics applications.

Result and discussion

The demonstration of the model was examined systematically by applying the key classification measures in the form of accuracy, recall, F1-score, and precision. These measures provide a comprehensive assessment of the model's ability to identify plant stress during its early stages. Accuracy measures the proportion of correctly classified instances out of the total instances and thus presents an overall idea of the model's performance. Accuracy computes the number of correctly labeled positive instances as a proportion of all cases estimated to be positive and hence analyzes how well the model performs in avoiding false positives. Recall, which is often referred to as sensitivity, analyzes the proportion of true positive cases the model identifies accurately, and it is essential to detect plant stress without missing any instances.

Conclusion

The comparison of KNN, Decision Tree, and Random Forest classifiers in the context of early plant stress detection emphasizes their efficiency in boosting classification performance. The result suggests that Random Forest, presents a valid and generalizable approach to the early detection of plant stress with relevance to timely intervention and improved crop management. Follow-up research can examine deep learning-based models for further improvement in the accuracy of classification and real-time detection.

Keywords: Classification metrics, Plant stress detection, machine learning