DOI: https://doi.org/10.52756/ijerr.2024.v46.014 *Int. J. Exp. Res. Rev.*, Vol. 46: 177-190 (2024)

Original Article | Peer Reviewed | Open Access

International Journal of Experimental Research and Review (IJERR)
© Copyright by International Academic Publishing House (IAPH)
ISSN: 2455-4855 (Online)
www.iaph.in

Automated Machine Learning Classification Framework to Predict Crop Yield and Detect Pest Patterns

Gopi R[1]*, Tamil Selvi M[2], Saranraj G[3], Nagaraj P[4], Parthiban K[5] and Ranjith Kumar A[6]

[1]Faculty of Computer Science & Engineering, Dhanalakshmi Srinivasan Engineering College, Perambalur–621212, Tamil Nadu, India; [2]Faculty of Computer Science and Engineering, Roever Engineering College, Perambalur-621212, Tamil Nadu, India; [3]Faculty of Artificial Intelligence and Data Science, Dhanalakshmi Srinivasan College of Engineering, Coimbatore 105, Tamil Nadu, India; [4]Faculty of Artificial Intelligence and Machine Learning, K. Ramakrishnan College of Engineering, Samayapuram - 621112, Tamil Nadu, India; [5]Faculty of Computer Science and Engineering, Dhanalakshmi Srinivasan College of Engineering, Coimbatore - 641105, Tamil Nadu, India; [6]Faculty of Computer Science and Engineering, Lovely Professional University, Phagwara, Punjab, India

E-mail/Orcid Id:

GR, gopi.r@dsengg.ac.in, https://orcid.org/0000-0003-4957-1843; *TSM*, tamilnaveena@gmail.com, https://orcid.org/0009-0003-4642-2807; *SG*, saranrajdsce@gmail.com, https://orcid.org/0009-0000-6412-115X; *NP*, pnagaraj.me@gmail.com, https://orcid.org/0009-0007-9438-1973; *PK*, janparthiban@gmail.com, https://orcid.org/0009-0007-2385-5529; *RKA*, ranjithdr.kumar@gmail.com, https://orcid.org/0000-0003-4383-9212

Article History:

Received: 01st Jun., 2024
Accepted: 18th Dec., 2024
Published: 30th Dec., 2024

Keywords:
Machine Learning, Plant Disease, Classification, Crop yield, Reliability

How to cite this Article:
Gopi R, Tamil Selvi M, Saranraj G, Nagaraj P, Parthiban K and Ranjith Kumar A (2024). Automated Machine Learning Classification Framework to Predict Crop Yield and Detect Pest Patterns. *International Journal of Experimental Research and Review*, *46*, 177-190.
DOI: https://doi.org/10.52756/ijerr.2024.v46.014

Abstract: Plant disease identification is crucial to food security and agricultural product availability. Traditional disease diagnosis can be tedious, annoying, and inaccurate. The investigation examines how modern machine learning algorithms might improve plant disease diagnostics for efficacy and precision. Despite this, machine learning faces many obstacles, including model training, processing costs, and rising demand for large data sets. This study proposes a novel method called Automated Machine Learning Classification Framework (AMLCF) to predict crop yield and detect pest patterns. This framework simplifies model selection, hyperparameter adjustment, and feature engineering for non-experts. The amount of time and computational resources needed have additionally been greatly reduced. The suggested AMLCF is evaluated on different unique agricultural datasets to validate its plant disease detection versatility. Our extensive simulation analysis found that AMLCF exceeds existing machine learning methods in speed, accuracy, and usability. AMLCF's detailed demonstration shows this; besides predicting plant illnesses, this system can predict crop yield and detect pests. Those findings suggest AMLCF could transform farming. Better plant health monitoring, early disease identification, and farmer selection could be achieved. The experimental results show that the proposed AMLCF model increases the accuracy ratio by 92.6%, computational efficiency analysis by 97.4%, versatility analysis by 98.3%, user accessibility ratio by 99.1%, and crop health tracking analysis by 94.8% compared to other existing models.

Introduction

Machine learning has revolutionized plant disease detection and these algorithms improve diagnostic efficiency and accuracy (Gong, 2021; Kaur, 2023; Rao et al., 2024; Prasad and Agniraj, 2024). The first issue is that plant diseases are diverse and unpredictable, making them hard to control. Plant symptoms are diverse, making it difficult for models to generalize across contexts (Ahmed and Ganapathy, 2021; Chaudhary et al., 2023). Plant species, infection stage, and weather can affect symptom severity and these models struggle to prove generalisability (Huang et al., 2021). Training requires big, diverse datasets, which can be difficult to get due to unpredictability (Wang, 2023). Many plant diseases share

***Corresponding Author:** gopi.r@dsengg.ac.in

symptoms (Yang, 2020). This increases false positives and negatives and makes discrimination harder, along with data quality, soabelling is crucial (Ahmad et al., 2022). Machine learning models need high-quality, well-annotated datasets (Pei and Wang, 2022).

Real-world agricultural applications of machine learning systems must understand the technology's limits (Yang et al., 2023). Portable or field-based computers have storage and processing limitations (Rong, 2021). Because farmers sometimes lack technical know-how, these devices must be straightforward (Shuliar et al., 2023). Models may be hard to adapt to unique geographies and environmental variables, affecting performance and illness development (Barakina et al., 2021).

Quality and quantity remain issues; gathering large datasets with suitable tags is expensive and complex (Fitria, 2021). Poorly trained or chaotic models may not generalize or be accurate. Model training is complicated and constrained by plant disease variability by species and habitat. Plant diseases vary significantly, making generalizations problematic (Yanru, 2021). Due to memory and processing power requirements, deep learning may be too expensive for field applications. Integrating the idea into real agricultural systems needs many steps, including making it farmer-friendly and responding to field circumstances. These difficulties must be resolved quickly to improve machine learning-based plant disease diagnosis.

The main contributions of the study are-

#Designing the AMLCF to employ contemporary machine learning to quickly and reliably diagnose plant diseases.

#The proposed system performs model selection, tweaking, and feature engineering to solve machine learning problems. Complex models and high processing costs are hurdles.

#The experimental results have been performed, and the suggested AMLCF model increases the accuracy ratio, computational efficiency analysis, versatility analysis, user accessibility ratio and crop health tracking analysis compared to other existing models.

#This section of the written research paper will conclude with the following outline: Section 2 focuses on modern machine-learning methods for precise plant disease identification. Section 3 presents the AMLCF. Section 4 offers a comprehensive appraisal, covering all aspects, such as effects and drawing connections to previous endeavours. Section 5 displays the results.

Related works

This examination delves into many innovative ways that can be used to improve educational outcomes through the integration of AI. This study aims to analyze AI-driven systems for online art classes, online basic education assessments, online preschool AI curricula, and online English classes to improve learning results and student engagement.

Sun et al., 2021 developed an AI-enhanced online English teaching system that employs deep learning and decision tree algorithms to deliver tailored lessons to each student. According to the test findings, the method significantly enhances students' learning capacity and the content's practicality.

An article by He and Sun (2021) presents AI-CATM, an artificial intelligence-driven computer-assisted teaching model designed to collaborate with teachers in developing personalized art lessons tailored to the unique needs of each student. This method substantially enhances study depth, student involvement, and creative design compared to traditional methods.

The study conducted by Li and Su (2020) employed EM-EW and grey clustering analysis to assess AI-integrated online basic education training. The model improves online teaching and provides insights for AI-driven basic education with its practical methods.

The article by Sun (2021) integrates "5G+AI" and holographic technologies into online spoken education for interactive learning using gesture detection. The strategy improves students' comprehension, critical thinking, and involvement over traditional methods.

The research by Su and Zhong (2022) uses problem-based learning (P-L) and social robots to create an artificial intelligence (AI) curriculum for preschools that is based on goals, content, methodologies, and assessment. By focussing on problem-based learning in particular, the method enhances AI literacy and the efficacy of learning in young children.

Nabarun Dawn et al. (2023) suggested that Artificial Intelligence, Machine Learning and the Internet of Things (IoT) revolutionize Agriculture. New tools, difficulties, and the potential future of artificial intelligence (AI) in agriculture are explored in depth in this study. In recent years, expert systems (ES), the IoT, and Artificial Neural Network (ANN) models may unleash AI's full potential in the agricultural sector.

Hemanta Gogoi et al. (2023) proposed the avian diversity in the paddy field ecosystem surrounding the Assam university campus in Silchar during the rainy season. There is a significant correlation between the amount of wasted rice and rice yields, and birds often eat newly planted rice seeds. Insect pests abound in paddy fields, which are a semi-aquatic habitat. Waterfowl, wading birds, shorebirds, and other species of waterbirds use these areas as staging grounds for migration and feeding grounds. We present 95 bird species, representing 37 families and 14 orders, from this pilot research carried out in the paddy field environment around the Silchar campus of Assam University.

In (Ilham Amani Rozaini et al., 2023) recommended the bilateral teleoperation with a shared design of master and slave devices for robotic excavators in agricultural applications. The operational complexity is greatly affected by the dynamic used in this teleoperation system, which is a master device that controls the activities of an agent device. Therefore, this work aims to improve the master-slave algorithm for teleoperation applications that depend on regulating the motions of robot arms. Although the master and slave devices are different in size, they are structurally similar. Ensing the understandable kinematic model that connects these parts is important for making the robot easy to manage and user-friendly. The robot arm's forward kinematics, computed using Denavit-Hartenberg parameters and transformation matrices, must be used to calculate the end effector movement and location.

Abderraouf Amrani et al. (2024) discussed the multi-task learning model for agricultural pest detection from crop-plant imagery. This approach uses a joint loss function, which merges a classification loss with a tailored size loss. An image containing an aphid may be located using the classification component, and its size can be approximated using the customized size loss function. The latter is personalized-made for a more precise size estimate to smooth out differences between predicted and measured ground truth sizes. This model is built with a ResNet18 backbone, which makes it very flexible and resilient.

Kariyanna and Sowjanya (2024) deliberated on theunravellingf use of artificial intelligence in managing insect pests. Using sophisticated algorithms, AI provides a game-changing method for analyzing complex data patterns gathered from many sources, such as sensors and pictures. Reducing the need for random pesticide applications and making the most of interventions allows for precise identification of pests, early diagnosis of problems, and predictive modelling, all of which improve pest management decision-making.

Philipp Batz et al. (2023) presented the potential of image recognition and artificial intelligence for aphid pest monitoring. Using aphids as an example, the author shows how future technological advances in systematic monitoring of insect pests, automated individual identification, and intelligent forecasting models might benefit from case data. Using aphids as an example, the author demonstrates how recent breakthroughs in image recognition technology can automate the process of identifying specific individuals in static images, opening the door to the possibility of systematic monitoring of insect pests.

Sangyeon Lee and Choa Mun Yun (2023) introduced sequential environmental data based on predicting risks of crop pests and diseases using deep learning models. The author demonstrated the model's ability to forecast the risk score of agricultural diseases and pests using large-scale public data on strawberry, pepper, grape, tomato, and paprika crops. Its average AUROC was 0.917, indicating strong predictive ability with these predictions, one may aid in preventing pests or post-processing. Many different facilities and crops may benefit from this learning framework and model for predicting crop diseases based on environmental data.

Sheela (2023) suggested Crop Yield Improvement with Weeds, Pest and Disease Detection. The data augmentation procedure is performed since Deep Learning performs better with bigger data sets. Various DCNN architectures were used to construct the neural model, and their accuracy and performance were used to interpret the models using hyperparameter-searching, InceptionV3, DenseNet201, Mobilenet, VGG16, and hyperparameter-tuning on data from agricultural picture sources. When comparing the end value, the tweaked InceptionV3 model performed 87.85% better. In contrast, Mobilenet and VGG16 got 91.85% and 78.71% accuracy, respectively. The DenseNet model outperformed the Hyperparameter Search with a 99.62% accuracy rate.

In Md. Akkas Ali et al. (2023) proposed sound analytics in large agricultural fields for pest detection systems using high-performance-oriented AI-enabled IoT. To denoise the pest sound, remove spectral leakage, convert overlapping to non-overlapping frames, convert the time to the frequency domain, determine the frequency spectrum, detect the sinusoidal frequency and internal component, and extract the MFCCs feature, the proposed method utilized audio pre-processing techniques in sound analytics, such as HPF, Hann window, hop window, FFT, DFT, STFT and the MFCC algorithm.

In Rashmi Priya Sharma et al. (2023) recommended the Internet of Farm Things-based prediction for crop pest infestation using an optimized fuzzy inference system (IoFT-FIS). Using these meteorological characteristics, the knowledge base of the proposed fuzzy inference system is constructed. The multi-objective evolutionary algorithm uses fuzzy criteria to determine an appropriate cropping window and breeding circumstances with low pest populations. Using an IEEE 802.15.4 wireless IoT sensor network monitoring infrastructure in medium grass vegetation, this proposal finds crop-sowing windows based on fuzzy logic with optimum crop production and lowest insect development. Crops of sugarcane and rice are the subjects of current experiments. This study's experimental setting was a farm near Gwalior, Madhya Pradesh, India. The field-deployed wireless sensor network gathered soil moisture, rainfall, temperature, and more data.

Galiya Anarbekova et al. (2024) discussed the fine-tuning of artificial neural networks to predict pest numbers in grain crops. The transformer reliably shows better prediction accuracy in terms of mean squared error. The effect of several training hyperparameters on predicted accuracy, including batch size and epochs, is also shown in this paper. Season 2's unusual reactions highlight the impact of certain characteristics on model performance, which is an intriguing development. This study adds to the growing knowledge of optimizing ANNs for precise insect prediction in grain crops, which helps create more effective and efficient pest management methods. It is also well-suited for real-world applications due to the transformer model's persistent dominance.

Figure 1. Block diagram for Data Collection.

Materials and Methods

The suggested approach uses an AMLCF, or Automated Machine Learning Classification Framework, to improve plant disease detection. AMLCF's use of advanced machine-learning techniques makes it easier for the system to capture images and the subsequent categorizing of diseases. The major goal is to increase agricultural production through improved efficiency and accuracy of plant disease detection, which helps crop health management. The suggested technique aims at a rapid, accurate diagnosis of plant diseases that strive for power with ease.

The suggested technique presents the AMLCF to forecast agricultural yields and identify pest trends accurately. This framework incorporates state-of-the-art machine learning algorithms. It aims to improve adaptation across agricultural datasets by focusing on automated feature selection and model improvement. Computationally efficient, scalable, and easily accessible, the framework is designed to make it easy for farmers and other stakeholders in the agricultural sector to deploy. The AMLCF offers a centralized platform for tracking pest activity and crop health, promoting proactive management and well-informed decisions in precision agriculture. To promote sustainable agricultural methods, these contributions tackle important issues in contemporary agriculture, such as optimizing resources and controlling pests promptly.

Contribution 1: Design of AMLCF

The following are three stages the recommended AMLCF goes through to improve detection efficiency greatly: Auto-preprocessing, segmentation, and feature extraction. This simplified process reduces computation resources and time needed for plant disease detection. Moreover, state-of-the-art machine learning approaches used in this framework boost classification precision, resulting in reliable and accurate identification of various plant diseases. Improved crop health management follows, as well as timely intervention actions.

Figure 1 shows the data-gathering procedure used to identify plant diseases within an AMLCF framework. The initial step involves capturing raw pictures of plants using cameras or sensors, which is called image acquisition. Each picture is then annotated to show

DOI: https://doi.org/10.52756/ijerr.2024.v46.014

relevant information, such as the kind of ailing plant, after which it is stored in an annotated dataset.

Then, these photos are pre-processed during the image processing phase to be better enhanced to find patterns. For instance, this stage can utilize noise reduction, contrast adjustments or segmentation methods to separate areas of interest depicted in the photos from each other. Next comes feature extraction, where certain attributes or features are recognized and extracted from the processed pictures. These attributes could be texture, colour and form, among others. These characteristic features are necessary for differentiating between different illnesses in plants. Lastly is classification, where machine learning methods are employed to sort the images into different illness categories based on the retrieved attributes. The model trained on the identified dataset to make illness predictions in fresh, unlabelled photos. Improved plant health surveillance and early disease diagnosis are made possible by this simplified technique, which is made possible by AMLCF, which increases the effectiveness and reliability of plant disease identification.

$$s_1 o^{r(a+d)} + o_2 r^{-r(a+d)} = o_3 r^{r(a)} + o_4 r^{-r(a)} \quad (1)$$

The Equation 1 may be connected to the suggested Automated Machine Learning Classification Framework s_1 as follows factors $(a+d)$ or elements that are tweaked o_2 throughout $-r,$ the procedure of machine learning is given by $a+d$ and $r(a)$, but the scores or indices modified to maximize the reliability of the model are denoted by $r^{-r(a)}$ and o_4.

$$\frac{v_3}{v_2} g^{r(a+d)+r(a)} + \frac{v_4}{v_2} h^{r(a+d)-r(a)} - \frac{v_1}{v_2} f^{2r(a+d)} \quad (2)$$

The equation 2 could be understood as $\frac{v_3}{v_2}$ representing various plant health indicators $\frac{v_4}{v_2}$, and g and h representing features $\frac{v_1}{v_2}$ extracted $a+d$ from agricultural datasets. Possible representations $r(a)$ of the model's sensitivity to changes in data or circumstances may be found in the exponents of $(a+d)$ and $-r(a)$.

This study's simulation environment was built using MATLAB Simulink, a powerful tool for modelling, simulating, and evaluating dynamic systems. The deciding factors were the environment's adaptability in incorporating machine learning models for prediction tasks and its capacity to manage massive agricultural information. Using past data on crop yields, patterns of pest infestation, and weather variables like temperature, humidity, and rainfall made the simulation seem like real-life agricultural circumstances. Because of the system's adaptability, seasonality and other regional considerations might be dynamically important. This study verified the suggested framework's accuracy and resilience in this controlled setting, which minimized the time and money needed for actual field trials.

Contribution 2: Implementation of Machine Learning

Model selection, hyperparameter tweaking, and feature engineering are just a few of the frequent

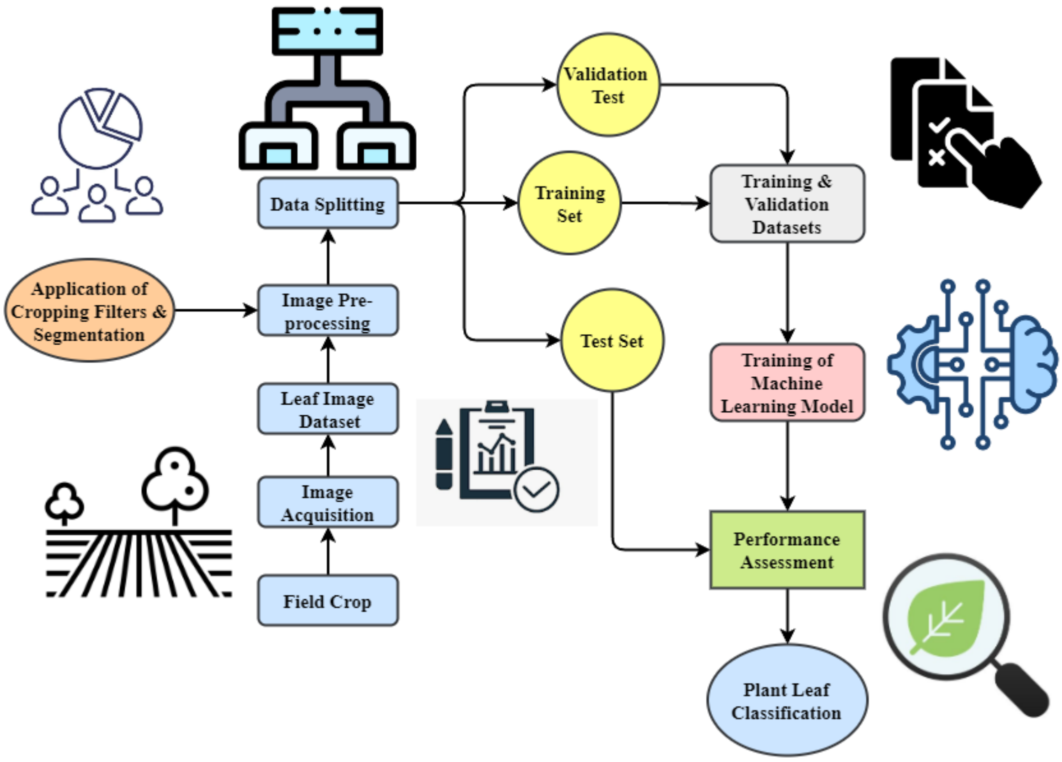

Figure 2. Process of Automated Machine Learning for Plant Disease deduction.

DOI: https://doi.org/10.52756/ijerr.2024.v46.014

machine-learning difficulties the AMLCF automates. The barrier to entry is lowered, and complex machine learning methods are accessible to non-experts because of this simplicity. More people will use and put the framework to use in the field because of how easy it is to use, even for those without much technical training, which is great for the agricultural industry.

The whole procedure of using machine learning technology to identify plant diseases is shown in Figure 2. The first step is to get photos from field crops, specifically from leaves, using either cameras or sensors. You may find these pictures in the Leaf Image Database. Image pre-processing improves the quality of the photos by using cropping filters and separation algorithms, which then isolate the leaf region. This phase is critical for precise disease identification because it zeroes in on the leaf and eliminates extraneous background data. Data Splitting is the next step after pre-processing the photos; it involves dividing the dataset into three parts: training, validation, and testing. This guarantees accurate model validation and training as well as reliable performance evaluation.

To train a machine learning model, one uses training and validation datasets. Here, the model is trained to recognize characteristics linked to different plant diseases in the training set, and its accuracy is adjusted with the help of the validation set. The trained model's performance is evaluated using the Test Set. In this stage, the model's accuracy in disease classification is tested with fresh, previously unseen photos of plants. Lastly, Plant Leaf Classification employs the performance-assessed model, which enables precise disease identification for efficient plant surveillance and early disease treatment. To improve the accuracy and efficiency of plant disease diagnoses, this AutoML approach simplifies the workflow.

$$\sigma_3\left(\frac{\vartheta_1}{\sqrt{1+\beta}}+\frac{\vartheta_2}{\sqrt{1-\beta}}\right)=\sigma_1\left(\frac{\vartheta_1}{\sqrt{1+\beta}}-\frac{\vartheta_2}{\sqrt{1-\beta}}\right)=0 \tag{3}$$

This is equation 3, seen as a balanced model inside (AMLCF), where various parameters (θ_1, θ_2) are standardized and changed according to a factor (β) to attempt to reach an ideal state. Machine learning relies on this equilibrium to achieve high accuracy and efficiency; it improves performance in agricultural applications like plant disease detection by ensuring features are appropriately scaled and weighted.

$$N(F)=\frac{1}{g}\sqrt{h-rl}/\partial(r-jf)=\tau\int_0^{\ni} i^{-hs+\frac{r}{fr}} \tag{4}$$

Equation 4 agreement with the suggested Automated Machine Learning Classification Framework $N(F)$. Accuracy of disease identification analysis, such as $\frac{1}{g}\sqrt{h-rl}$, and $\partial(r-jf)$ reflect various model and data attributes that impact τ the learning process, whereas $i^{-hs+\frac{r}{fr}}$ denotes the model's performance metrics. Simplifies difficult processes $\ni$ for enhanced plant disease detection $\frac{r}{fr}$ by balancing computing efficiency and accuracy.

Contribution 3: Demonstrate Real-World Benefits and Versatility

The AMLCF's flexibility in effectively identifying many plant diseases is shown by testing it on diverse agricultural datasets, showcasing its real-world application. The framework has other uses beyond only diagnosing diseases; it may identify pests and monitor the condition of crops. The impressive results in many settings highlight its ability to transform farming methods, leading to better yields and more efficient operations. The AMLCF's versatility makes it a useful tool in many agricultural settings.

In Figure 3, machine learning for plant disease categorization is displayed. The Dataset has to be expanded as the first step so that the model can accommodate more data. To enhance the images further, Image Filtering is done on the dataset, removing noise and other irrelevant information. Attributes Selection is the next stage in illness classification that comes after image filtering to determine which pixels in an image are important. This processing stage determines what leaf traits indicate what diseases; hence, it's crucial for the model's reliability. After this, datasets are separated into two parts, namely, training and test sets. Training data trains a machine learning model used by models, while test data assesses its performance in terms of accuracy and generalizability. This ensures that our model can generalize well even on unknown samples.

The above approach gives us a Machine Learning algorithm that trains on our training data and finds features associated with various diseases of plants. The accuracy and efficiency of such a trained model are then determined when applied to test data. A trained model is applied to predict three diseases: brown spots, leaf bacteria blight, and leaf moulds. Like any other machine-learning technique, this allows future detection of three crop sicknesses: brown spots, leaf bacteria blight, and leaf Molds.

$$\sum_{k=1}^{s}\alpha_l+d_{l+r}-sy_{t+k}>0,\propto_q+MR \tag{5}$$

A representation of the classification model's

DOI: https://doi.org/10.52756/ijerr.2024.v46.014

threshold-based decision-making $k = 1$ process when Equation 5 is satisfied. The parameters or weights that have been learnt, α_l and d_{l+r}, aspects or data elements from the dataset, $0, \propto_q$ and sy_{t+k}, computational efficiency analysis and a margin or regularization term, MR, are all presented this time.

$$G_{iR}\big(u_R(x)\big) = G_i\big(u_R(x)\big) + a.e.x \in \tau, \sigma\, i = 1, \ldots, N \quad (6)$$

variables $a.e.x$ and $\sigma\, i$, guarantee the implementation of customized solutions N for different agricultural datasets.

The AMLCF, a computerized machine learning classification framework, is shown in Figure 4. The data are from the Agriculture Crop Images Kaggle Dataset. The Crop images dataset includes 40 or more images of every agricultural crop, including sugarcane, rice, jute, wheat, and maize. There are more than 160 enhanced

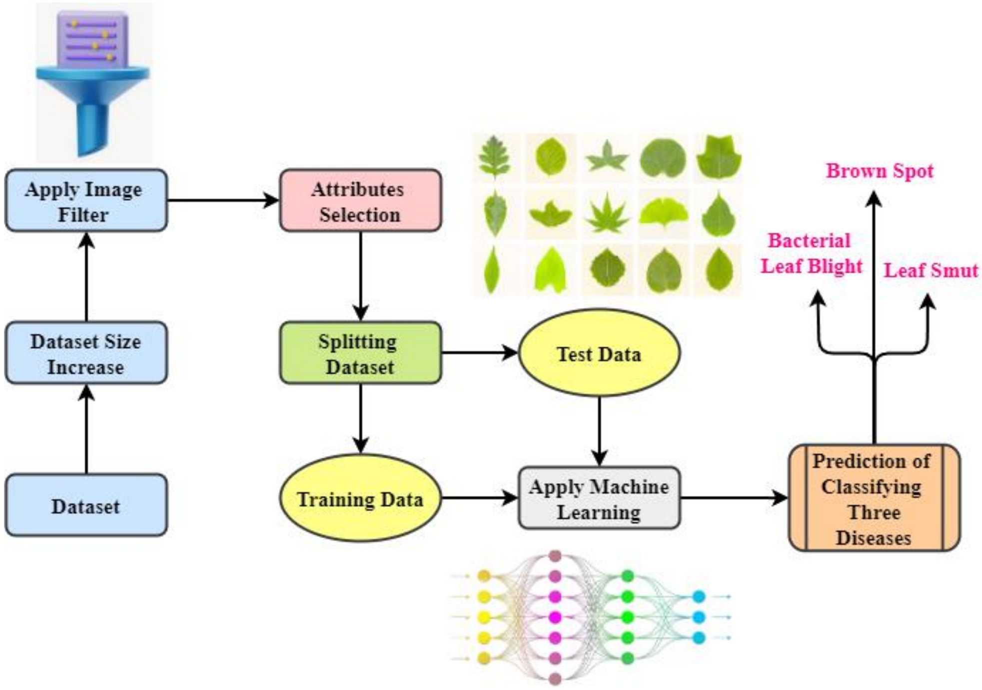

Figure 3. Plant disease classification using machine learning.

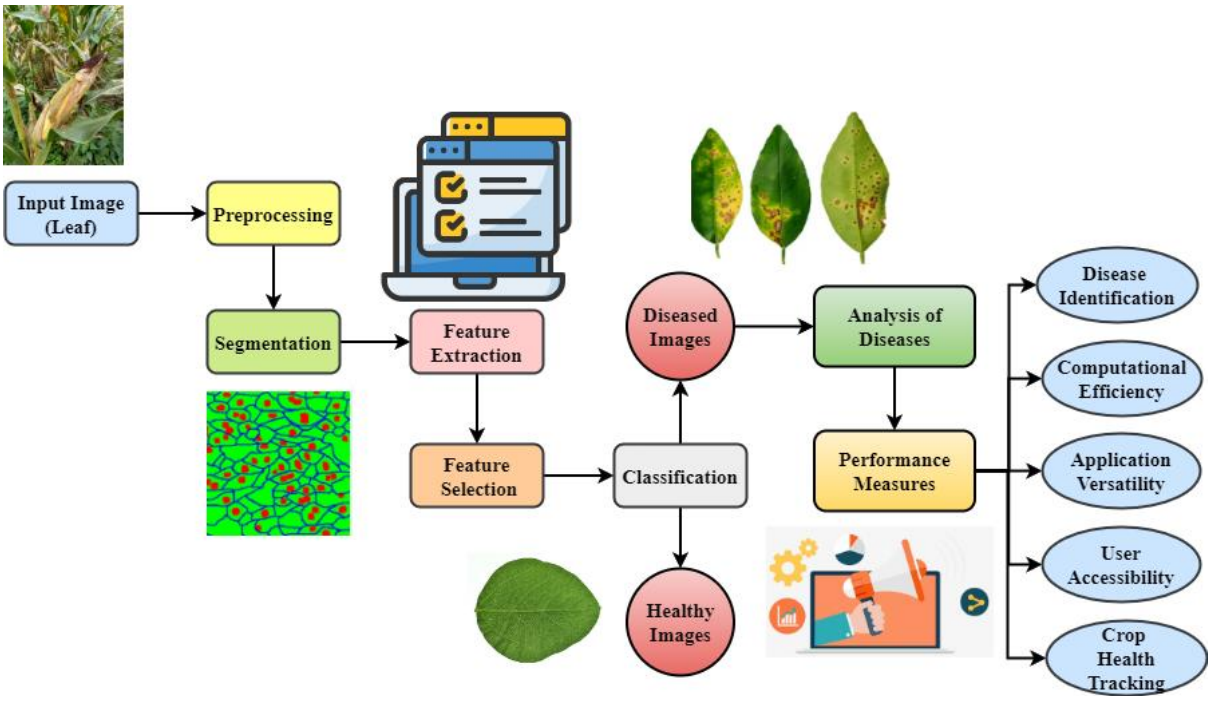

Figure 4. The overall structure of AMLCF.

The Equation 6 may be linked to the offered G_{iR}. To improve the algorithm for improved performance $u_R(x)$, this change is like application versatility analysis hyperparameter tweaking G_i and feature engineering. The model's thresholds or limitations are represented by the crop images in each class in the dataset (kag2). Among the augmentation features are the ability to rotate, flip horizontally, and shift vertically. Images of each class show them at various stages of their life cycle, from above to below and from aerial to ground level.

DOI: https://doi.org/10.52756/ijerr.2024.v46.014

Improving the accuracy of crop image classification for 5 agricultural crop types: wheat, rice, sugarcane, maize, and jute.

First, cameras or sensors are used to take pictures of the plant's leaves, known as the Input Image Leaf. After that, it goes on to pre-processing, where contrast adjustment and noise reduction are used to improve the photos. After the pre-processing, the segmentation procedure is carried out to separate the leaf area from the backdrop. This ensures that the emphasis is on the correct area. Next, Feature Extraction finds the leaves' unique characteristics, such as their form, colour and texture, important for differentiating between healthy and sick leaves. The next step, Classification, uses the extracted data to classify the leaves as healthy or sick using machine learning algorithms.

After completing the classification, the system creates Infected and Healthy Images to validate the findings visually. The Analysis of Disease section delves further into the recognized disorders to better understand them and record our findings. Reliable disease detection is achieved using Performance Measures to evaluate the classification model's accuracy and efficiency. The computational efficiency of the framework is maximized by reducing processing time and resource consumption. To facilitate successful interaction between the system and farmers and agricultural specialists, User Accessibility guarantees that the system is straightforward to use. Crop health tracking is another area where AMLCF has found use. This technology allows for the constant monitoring and control of plant health, which helps with improved farming methods and the early diagnosis of diseases.

$$I_a = [0, a] \cap T = \{t \in T : 0 \leq t \leq a\} = [0, a]T \tag{7}$$

The Equation 6 could help reduce computational resources and improve model efficiency I_a and accuracy $[0, a]$ by selecting time-specific features or data points $\cap$ T for model training. Improved disease diagnosis alongside $t \in T$ User accessibility analysis is possible because of t ability to handle and evaluate a the temporal features $[0, a]$ of agricultural information T by concentrating on this subset.

$$f^{\Delta}(t) = \lim_{s \to t} \frac{f(\sigma(t)) - f(s)}{\sigma(t) - s} \quad i = 1, \ldots, N \tag{8}$$

This is the equation 8 function f^{Δ} about a timescale (t). By seeing $f(\sigma(t))$ as the process (f) of adaptive feature selection $\lim_{s \to t}$ and model tuning over time. The enhancer of the function $f(s)$ crop health tracking analysis as AMLCF changes and modifies its models using fresh data (t) demonstrates the system's capability for fluid adaptation $i = 1, \ldots, N$ and refining.

According to its simplified and automated methodology, the suggested AMLCF technique shows substantial gains in plant disease diagnosis. Beginning with image capture, the approach consists of many fine-tuned phases for accuracy and efficiency: pre-processing, segmentation, extracting features, and classification. Testing the framework on various agricultural datasets has shown that the approach outperforms conventional methods. In addition to detecting plant illnesses, the AMLCF helps monitor crop health and manage diseases early on. Intending to keep crops healthy, this method is useful for farmers and experts in agriculture as it guarantees efficient computing, high accuracy, and user accessibility.

Result and Discussion

Accuracy, computational efficiency, application adaptability, user accessibility, and crop health tracking are some metrics used to assess the efficacy of advanced machine learning algorithms in plant disease diagnosis. The MATLAB Simulink environment, a flexible tool for modelling and simulating dynamic systems, was used to perform the simulation tests. The capacity to include machine learning models with agricultural datasets and the simulator's reliability in handling time-series data were the deciding factors in its selection. A time step of 0.01 seconds and a simulation length of 10,000 seconds were set up for the experiment. The environmental elements included temperature (20-35°C) and humidity (40-60%), while the crop-specific variables were growth rate and insect infestation thresholds. These parameter sets were evaluated against available literature to guarantee realism and derived from experimental findings. The system allowed for a comprehensive evaluation of the efficacy of agricultural production prediction and pest detection by creating a controlled simulation environment that successfully mimicked real-world environments.

In Figure 5 above, controlling plant diseases efficiently requires this ability; CNNs and deep learning are two methods that have substantially improved diagnostic accuracy. This has been achieved by carefully examining patterns in botanical images. Despite their ability to efficiently distinguish between different diseases, the effectiveness of these systems is greatly reliant on the diversity and quality of the training data utilized. It is critical to obtain high-quality, adequately tagged datasets to train an accurate model that produces 92.6%. Furthermore, thorough model tweaking and

DOI: https://doi.org/10.52756/ijerr.2024.v46.014

validation are required to reduce the number of mistakes. AMLCFs can improve accuracy; these frameworks streamline model selection and hyperparameter tuning, making these procedures more user-friendly and effective. There has been a lot of improvement; however, there are still problems like data unpredictability and the strain on computer resources. Consequently, constant improvement and updating are necessary. Nevertheless, it is crucial to address these concerns to keep performance consistent across various agricultural situations, even while strong machine learning algorithms greatly improve disease identification accuracy.

learning models and CNNs generally need a lot of memory and processing power yet produce very accurate results. In particular, when working with massive datasets, the complicated computations needed to train these models can drive up the associated costs and increase the implementation duration by 97.4%. The development of AMLCFs is a response to these problems. These frameworks can make procedures like model selection and hyperparameter tuning easier. Further, these frameworks expedite model deployment while reducing the amount of computational labour needed. Using AMLCFs will make getting optimal

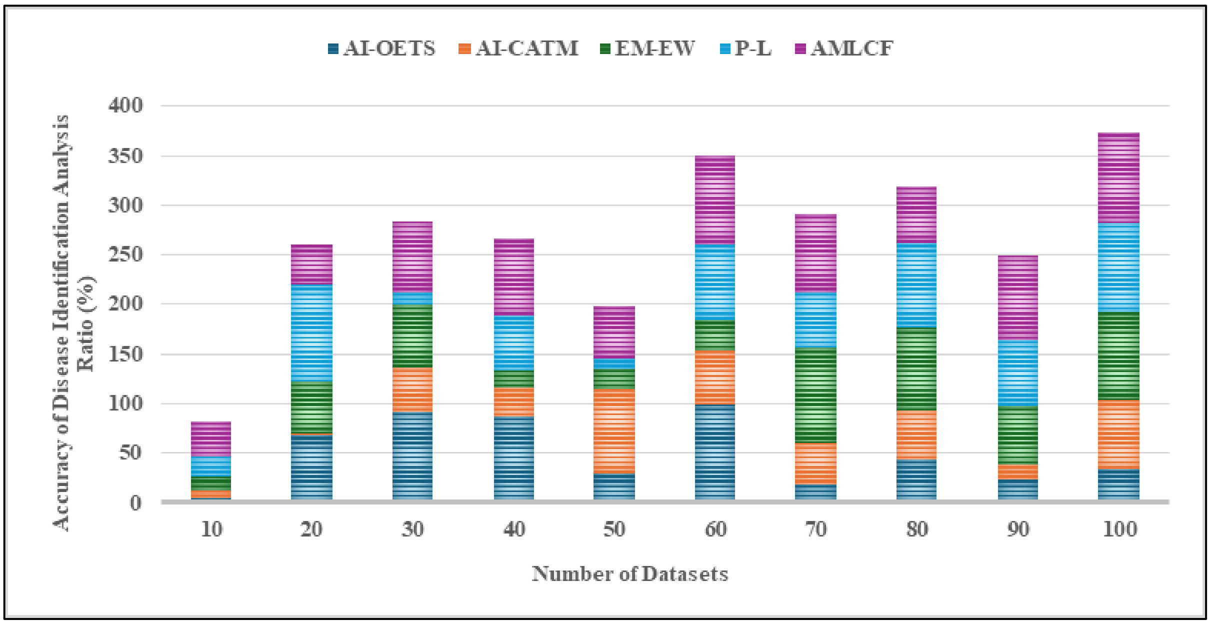

Figure 5. Accuracy of Disease Identification Analysis.

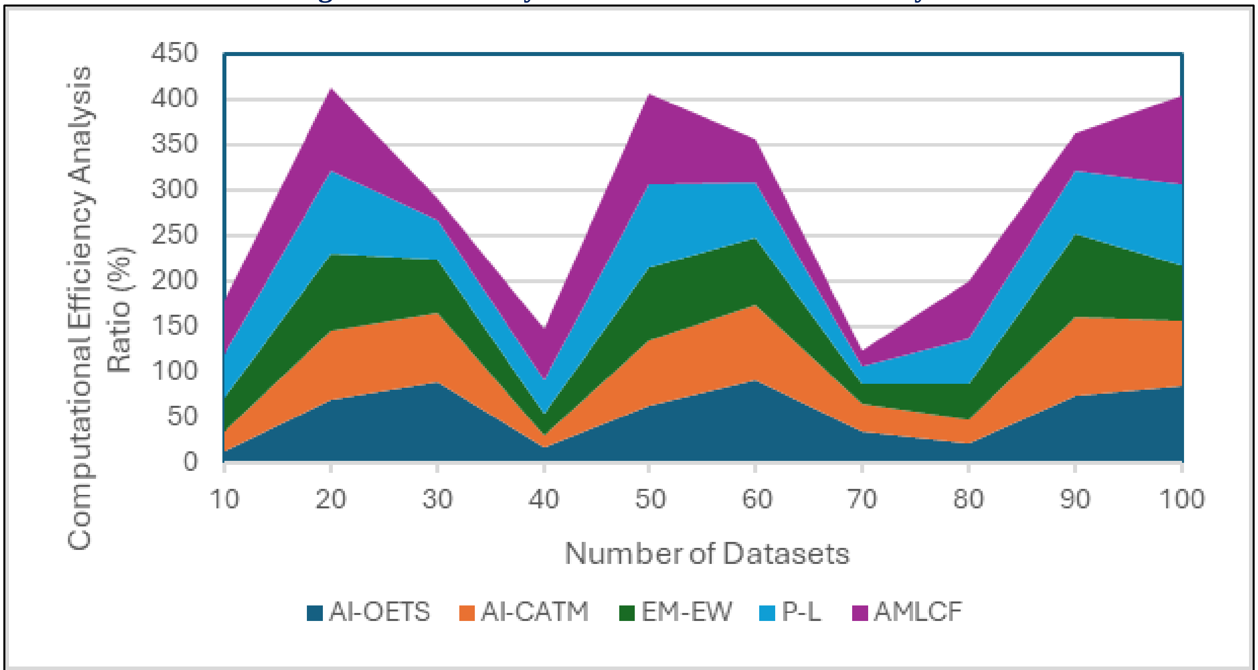

Figure 6. Computational Efficiency Analysis.

In Figure 6 above, analyzing the computational efficiency of potential methods for plant disease diagnostics is a crucial part of evaluating advanced machine learning algorithms. Techniques like deep

performance with fewer resources a lot easier. Simplifying and automating these processes will make this possible. Hardware efficiency and fine-tuned algorithms are prerequisites for capable control of

DOI: https://doi.org/10.52756/ijerr.2024.v46.014

resource utilization. This is because, despite advancements, deep learning algorithms can still retain very high computing demands. Finding an appropriate balance between computational economy and efficiency is crucial for real-world applications in agricultural settings, even when modern approaches guarantee a high degree of accuracy.

essential to consider the user's accessibility initially. Technical proficiency and specialized knowledge are frequently required to apply and interpret new approaches that offer high accuracy, such as CNNs. Because they streamline the otherwise convoluted procedures of model selection, hyperparameter tuning, and feature engineering, AMLCF can alleviate this problem. This

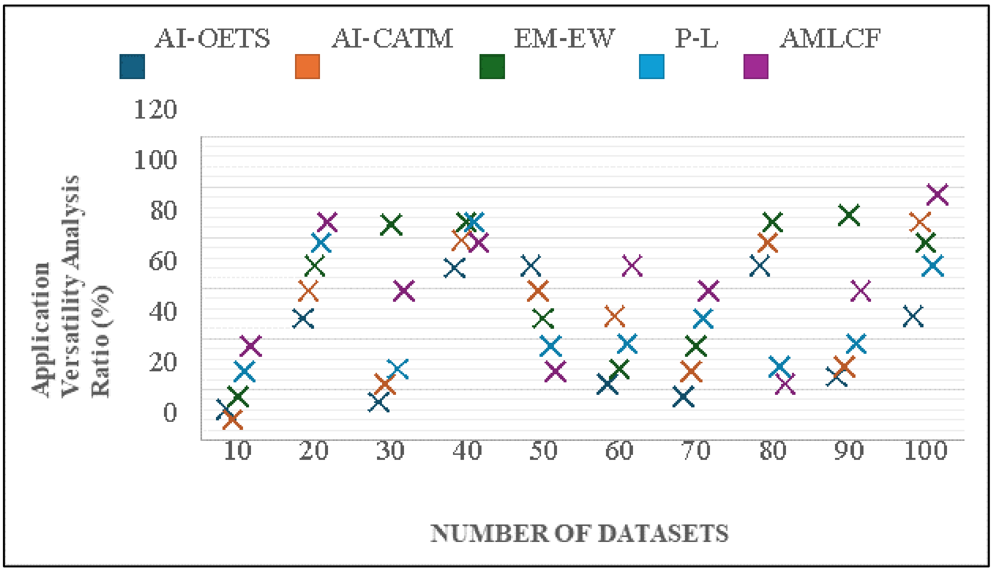

Figure 7. Application Versatility Analysis.

In Figure 7 above, the wide range of applications of modern machine learning algorithms in plant disease identification demonstrates their versatility. Methods like CNNs prove their adaptability by detecting numerous plant diseases in various environments and with different species. This proves that CNNs can discriminate between various plant diseases with high accuracy. AMLCFs enhance this adaptability by making model modification easier and facilitating rapid adaptation to varied datasets and illness types. In addition to helping with disease identification, these frameworks could be useful for other agricultural tasks, including estimating crop yields and preventing pest infestations. Because of the adaptability of modern machine learning algorithms, they may be used for a broad range of agricultural tasks, allowing for comprehensive plant health monitoring, which produces 98.3%. To ensure consistent performance across a wide number of applications and conditions, it is necessary to continually improve and validate it, even if it is adaptable. This is the sole method to guarantee that there will be ongoing development. By offering complete solutions to numerous problems related to plant disease control, these strategies may cause an unprecedented shift in agricultural operations. The methods' broad adoption demonstrates their promise and highlights their ability to revolutionize agricultural operations.

In Figure 8, when applying modern machine learning algorithms to the problem of plant disease diagnosis, it is makes these tactics more accessible to those who may not have a strong background in technology. By automating formerly manual procedures, AMLCFs lessen the demand for specialized skills and facilitate their incorporation into real-world agricultural contexts. This paves the way for their usage to spread, producing 99.1%. The availability of user-friendly support tools and interfaces is crucial to ensuring farmers and agricultural professionals can effectively utilize new technologies. Although some success has been had, further work is required to make the design and support even more user-friendly. Despite some success, this remains the case. By taking this route, we can make sure that innovative machine-learning methods are applied to farming in an approachable and beneficial way.

In Figure 9, advanced machine learning methods greatly improve the capacity to monitor crop health by allowing for the rapid and precise diagnosis of diseases. Methods like CNNs do an excellent job of analyzing plant images, which can be used to detect disease signs in plants at an early stage. This enables the proactive management of crop health. The method is further optimized with AMLCFs, and improved access to advanced disease diagnostics results from these frameworks' efforts streamline model deployment and reduce processing requirements. Farmers can get real-time alerts and information about possible problems by incorporating these techniques into crop health

DOI: https://doi.org/10.52756/ijerr.2024.v46.014

monitoring systems, which produce 94.8%. Because of this, farmers can now control and intervene with their crops at the perfect moment. Additionally, this preventative measure encourages better crop treatment decisions and resource allocation and improves the accuracy of illness diagnosis.

inevitably be obstacles that must be overcome. When considering all factors, crop health management benefits substantially from applying advanced machine learning algorithms. Table 1 shows the performance analysis.

Despite these advancements, complications like data variability and computational resource demands persist.

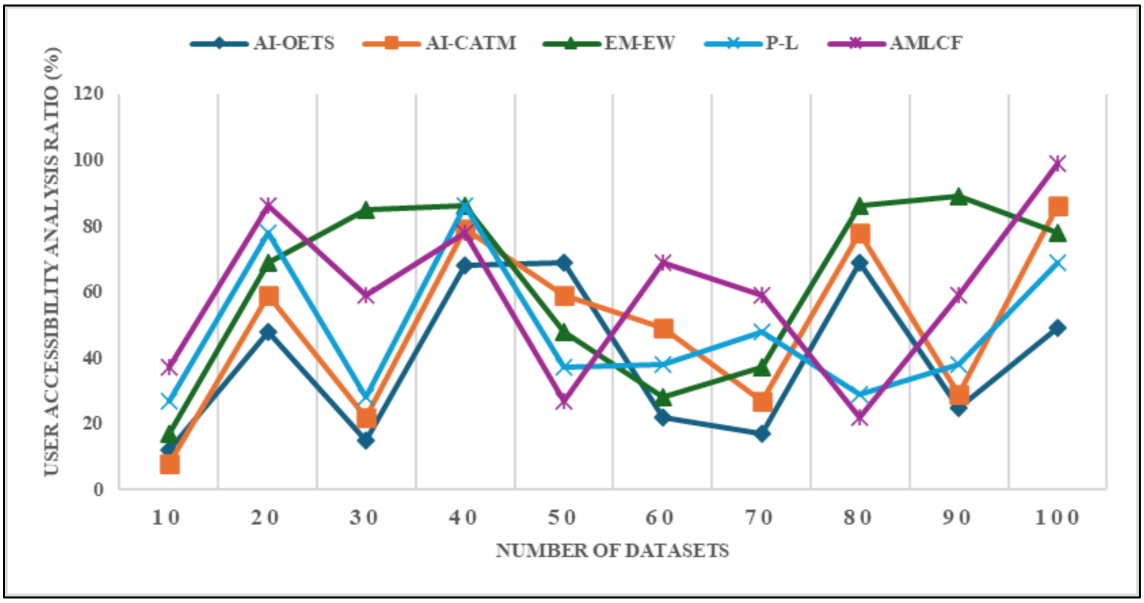

Figure 8. User Accessibility Analysis.

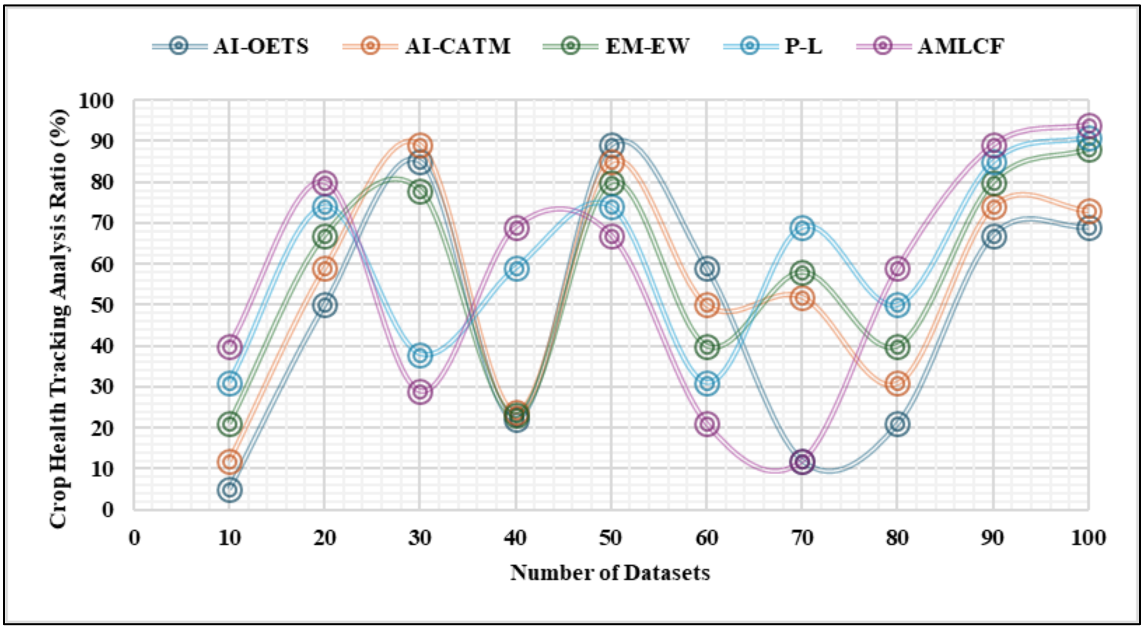

Figure 9. Crop Health Tracking Analysis.

Table 1. Performance Analysis.

Parameter	AMLCF	AI-OETS	AI-CATM	EM-EW	P-L
Accuracy Ratio (%)	92.6	85.4	88.2	83.7	84.9
Computational Efficiency (%)	97.4	89.3	91.1	87.6	88
Versatility Analysis (%)	98.3	90.2	92.5	89	91.3
User Accessibility Ratio (%)	99.1	92.4	94	91.2	92
Crop Health Tracking (%)	94.8	87.5	89.8	86.2	88

Ongoing issues, such as data quality and model flexibility, must be addressed if crop health tracking remains accurate and helpful across various agricultural contexts. Despite potential advantages, there will

This highlights the critical need for ongoing improvement to ensure efficient and consistent application in various agricultural contexts.

Conclusion

The following section of the paper describes how advanced machine-learning algorithms for plant disease identification can revolutionize the field. Traditional plant disease identification takes time and effort, reducing food safety and agricultural output. The research solves the issues completely. An AMLCF is developed to simplify feature engineering, model selection, and hyperparameter tuning. Advanced machine learning methods become more accessible and efficient. AMLCF exceeds standard accuracy, speed, and usability approaches when applied to agricultural datasets. This diagnostic paradigm helps agriculture anticipate crop yields, identify pests, and diagnose plant diseases. These findings suggest that AMLCF could revolutionize agriculture by improving crop health monitoring, early disease detection, and farmer decision-making. The fact that AMLCF can reduce computer expenses and simplify operations makes it suitable for modern agriculture. This improves sustainable and effective agriculture, and an AMLCF tool addresses crucial plant disease control challenges. The experimental results show that the proposed AMLCF model increases the accuracy ratio by 92.6%, computational efficiency analysis by 97.4%, versatility analysis by 98.3%, user accessibility ratio by 99.1%, and crop health tracking analysis by 94.8% compared to other existing models. Incorporating satellite images and data from IoT sensors for real-time monitoring, the framework may be enhanced to include more varied information from other geographies and crops in future research. Additionally, using advanced explainable AI approaches may make the model predictions more interpretable, which is great for farmers and agricultural professionals since it allows them to make better decisions.

Conflict of Interest

The authors declare that there is no conflict of interest.

References

Ahmad, S. F., Alam, M. M., Rahmat, M. K., Mubarik, M. S., & Hyder, S. I. (2022). Academic and administrative role of artificial intelligence in education. *Sustainability*, *14*(3), 1101. https://doi.org/10.3390/su14031101

Ahmed, A. A. A., & Ganapathy, A. (2021). Creation of automated content with embedded artificial intelligence: a study on learning management system for educational entrepreneurship. *Academy of Entrepreneurship Journal*, *27*(3), 1-10.

Ali, Md. A., Dhanaraj, R. K., & Nayyar, A. (2023). A high-performance-oriented AI-enabled IoT-based pest detection system using sound analytics in large agricultural field. *Microprocessors and Microsystems*, *103*, 104946. https://doi.org/10.1016/j.micpro.2023.104946

Amrani, A., Diepeveen, D., Murray, D., Jones, M. G. K., & Sohel, F. (2024). Multi-task learning model for agricultural pest detection from crop-plant imagery: A Bayesian approach. *Computers and Electronics in Agriculture*, *218*, 108719. https://doi.org/10.1016/j.compag.2024.108719

Anarbekova, G., Ruiz, L. G. B., Akanova, A., Sharipova, S., & Ospanova, N. (2024). Fine-Tuning Artificial Neural Networks to Predict Pest Numbers in Grain Crops: A Case Study in Kazakhstan. *Machine Learning and Knowledge Extraction*, *6*(2), 1154–1169. https://doi.org/10.3390/make6020054

Barakina, E. Y., Popova, A. V., Gorokhova, S. S., & Voskovskaya, A. S. (2021). Digital Technologies and Artificial Intelligence Technologies in Education. *European Journal of Contemporary Education*, *10*(2), 285-296. https://doi.org/10.13187/ejced.2021.2.285

Batz, P., Will, T., Thiel, S., Ziesche, T. M., & Joachim, C. (2023). From identification to forecasting: the potential of image recognition and artificial intelligence for aphid pest monitoring. *Frontiers in Plant Science*, *14*. https://doi.org/10.3389/fpls.2023.1150748

Chaudhary, Yashi, & Pathak, H. (2023). MCIP: Mining Crop Image Data On pyspark data frame Using Feature Selection and Cluster Based Techniques. *Int. J. Exp. Res. Rev.*, *34*(Special Vol.), 106-119. https://doi.org/10.52756/ijerr.2023.v34spl.011

Dawn, N., Ghosh, T., Ghosh, S., Saha, A., Mukherjee, P., Sarkar, S., Guha, S., & Sanyal, T. (2023). Implementation of Artificial Intelligence, Machine Learning, and Internet of Things (IoT) in revolutionizing Agriculture: A review on recent trends and challenges. *International Journal of Experimental Research and Review*, *30*, 190–218. https://doi.org/10.52756/ijerr.2023.v30.018

Fitria, T. N. (2021, December). Artificial intelligence (AI) in education: Using AI tools for teaching and learning process. *In Prosiding Seminar Nasional & Call for Paper STIE AAS*, *4*(1), 134-147. https://prosiding.stie-aas.ac.id/index.php/prosenas/article/view/106

Gogoi, H., Purkayastha, J., & Roychoudhury, S. (2023). Avian diversity in the paddy field ecosystem surrounding the Assam University campus in Silchar during the rainy season. *International Journal of*

Experimental Research and Review, 34(Special Vol.), 120–137. https://doi.org/10.52756/ijerr.2023.v34spl.012

Gong, Y. (2021). Application of virtual reality teaching method and artificial intelligence technology in digital media art creation. *Ecological Informatics, 63*, 101304. https://doi.org/10.1016/j.ecoinf.2021.101304

He, C., & Sun, B. (2021). Application of artificial intelligence technology in computer-aided art teaching. *Computer-Aided Design and Applications, 18*(S4), 118-129. https://doi.org/10.14733/cadaps.2021.S4.118-129

https://www.kaggle.com/datasets/aman2000jaiswal/agriculture-crop-images?resource=download

Huang, J., Saleh, S., & Liu, Y. (2021). A review on artificial intelligence in education. *Academic Journal of Interdisciplinary Studies, 10*(3). https://doi.org/10.36941/ajis-2021-0077

Kariyanna, B., & Sowjanya, M. (2024). Unravelling the use of artificial intelligence in management of insect pests. *Smart Agricultural Technology*, *8*, 100517. https://doi.org/10.1016/j.atech.2024.100517

Kaur, P. (2023). Performance and Accuracy Enhancement During Skin Disease Detection in Deep Learning. *Int. J. Exp. Res. Rev.*, *35*, 96-108. https://doi.org/10.52756/ijerr.2023.v35spl.009

Lee, S., & Yun, C. M. (2023). A deep learning model for predicting risks of crop pests and diseases from sequential environmental data. *Plant Methods, 19*(1). https://doi.org/10.1186/s13007-023-01122-x

Li, M., & Su, Y. (2020). Evaluation of online teaching quality of basic education based on artificial intelligence. *International Journal of Emerging Technologies in Learning (iJET), 15*(16), 147-161. https://doi.org/10.3991/ijet.v15i16.15937

Pei, Z., & Wang, Y. (2021). Analysis of computer-aided teaching management system for music appreciation course based on network resources. *Computer-Aided Design and Applications*, *19*(S1), 1-11. https://doi.org/10.14733/cadaps.2022.S1.1-11

Prasad, P., & Agniraj, S. (2024). Cross-Domain Adaptation Techniques for Robust Plant Disease Detection: A DANN-CORAL Hybrid Approach. *International Journal of Experimental Research and Review*, *42*, 68-84. https://doi.org/10.52756/ijerr.2024.v42.007

Rao, K., Devi, J., Anuradha, Y., G, K., Kumar, M., & Rao, M. S. (2024). Enhancing Liver Disease Detection and Management with Advanced Machine Learning Models. *International Journal of Experimental Research and Review*, *42*, 100-110. https://doi.org/10.52756/ijerr.2024.v42.009

Rong, L. (2021). Design of Ideological and Political Multimedia Network Teaching Resources Integration System Based on Wireless Network. *Scientific Programming, 2021*(1), 4293771. https://doi.org/10.1155/2021/4293771

Rozaini, I. A., Ahmad Zakey, N. E. N., Mohd Zaman, M. H., Ibrahim, M. F., Mustaza, S. M., & Mohamed Moubark, A. (2023). Bilateral Teleoperation with a Shared Design of Master and Slave Devices for Robotic Excavators in Agricultural Applications. *International Journal of Experimental Research and Review*, *35*, 119–127. https://doi.org/10.52756/ijerr.2023.v35spl.011

Sharma, R. P., Dharavath, R., & Edla, D. R. (2023). IoFT-FIS: Internet of farm things based prediction for crop pest infestation using optimized fuzzy inference system. *Internet of Things*, *21*, 100658. https://doi.org/10.1016/j.iot.2022.100658

Shuliar, V., Shkurko, V., Polukhtovych, T., Semeniako, Y., Shanaieva-Tsymbal, L., & Koltok, L. (2023). Using Artificial Intelligence in Education. BRAIN. *Broad Research in Artificial Intelligence and Neuroscience*, *14*(3), 516-529. https://doi.org/10.18662/brain/14.3/488

Su, J., & Zhong, Y. (2022). Artificial Intelligence (AI) in early childhood education: Curriculum design and future directions. *Computers and Education: Artificial Intelligence*, *3*, 100072. https://doi.org/10.1016/j.caeai.2022.100072

Sun, X. (2021). 5G joint artificial intelligence technology in the innovation and reform of university English education. *Wireless Communications and Mobile Computing*, *2021*(1), 4892064. https://doi.org/10.1155/2021/4892064

Sun, Z., Anbarasan, M., & Praveen Kumar, D. J. C. I. (2021). Design of online intelligent English teaching platform based on artificial intelligence techniques. *Computational Intelligence*, *37*(3), 1166-1180. https://doi.org/10.1111/coin.12351

Wang, Y. (2020). Analysis on the construction of ideological and political education system for college students based on mobile artificial intelligence terminal. *Soft Computing*, *24*(11), 8365-8375. https://doi.org/10.1007/s00500-020-04932-6

Yang, H., Shankar, A., & Velliangiri, S. (2022). Artificial intelligence-enabled interactive system modeling for teaching and learning based on cognitive web services. *International Journal of e-Collaboration (IJEC)*, *19*(2), 1-18. https://doi.org/10.4018/ijec.316655

Yang, R. (2020). Artificial intelligence-based strategies for improving the teaching effect of art major courses in colleges. *International Journal of Emerging Technologies in Learning (IJET)*, *15*(22), 146-160. https://doi.org/10.3991/ijet.v15i22.18199

YanRu, L. (2021). An artificial intelligence and machine vision-based evaluation of physical education teaching. *Journal of Intelligent & Fuzzy Systems*, *40*(2), 3559-3569. https://doi.org/10.3233/JIFS-189392

How to cite this Article:

Gopi R, Tamil Selvi M, Saranraj G, Nagaraj P, Parthiban K and Ranjith Kumar A (2024). Automated Machine Learning Classification Framework to Predict Crop Yield and Detect Pest Patterns. *International Journal of Experimental Research and Review*, *46*, 177-190.

DOI : https://doi.org/10.52756/ijerr.2024.v46.014