International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(15s) 2026 pp. 343-361
https://doi.org/10.70917/ijcisim-2026-4407
Article

HECYO-NET: HYBRID ENSEMBLE MOTH FLAME AND SHARK SMELL OPTIMIZATION MODEL FOR CROP MANAGEMENT

G. Alexandar Narkunam[1*], K. Kala[2]

[1*]Department of Computer Science, Alagappa University, Karaikudi - 630 003, India
[2]Department of Computer Science, Nachiappa Swamigal Arts and Science College, (Affiliated to Alagappa University), Koviloor- 630307, India
[*]Corresponding Author Email Id: narkunam2015@gmail.com
Co-author Email Id: kalakasinathan06@gmail.com

Abstract: Precision agriculture needs accurate, scalable and adaptive models to handle issues like noisy data, sensitivity to outliers, heterogeneity and generalizability of crop yield prediction. Conventional machine learning and deep learning (DL) algorithms tend to be unreliable because of environmental conditions and inconsistencies in large-scale data. To address these issues, this paper proposes the Hybrid Evolutionary Crop Yield Optimizer Network (HECYO-Net), which combines state-of-the-art preprocessing, feature engineering, and a new Hybrid Flame-Shark Evolutionary Optimizer (HFSEO). The framework uses K-Nearest Neighbours (KNN) based imputation with robust scaling as a preprocessing step, denoising autoencoders to reduce noise and extract latent features and ReLU-based ensemble construction trained with Root Mean Square Propagation (RMSProp). The optimisation is carried out by HFSEO, a hybrid of Moth Flame Optimization (MFO) and Shark Smell Optimization (SSO) to provide convergence and escape local minima. Deep Q-Networks (DQN) are also applied to make adaptive agricultural recommendations. Experimental results on a large dataset of 1,000,000 agricultural records demonstrate that HECYO-Net obtains the best performance, MAE = 0.09, MSE = 0.05, RMSE = 0.14, MAPE = 0.09, and nRMSE = 0.28, compared to CatBoost, XGBoost, LightGBM, CNN-RNN, and other ensemble baselines. The results demonstrate the importance of the model in delivering sustainable, precise and computationally effective crop yield prediction and decision support to the farmers.
Keywords: Precision agriculture, crop yield prediction, Hybrid Flame–Shark Evolutionary Optimizer, ensemble learning, denoising autoencoder, KNN imputation, reinforcement learning.

1. INTRODUCTION

The crop yield refers to the quantity of crops that have been harvested a product such as meat, wool, and milk, per unit area of land. Then the alternative method for evaluating the yield is the seed ratio. In India, crop yield in agriculture is generally globally below average due to the ages of major crops [1-2]. Although India is one of the leading agricultural producers, and sometimes the crop yield is very low compared to other nations such as the United States of America. For this crop yield data analysis, the data from various nations have to be collected, and then the analysis has to be done based on the crop yield [3]. The data set contains the fertilizers, pesticides, each crop's yield, annual rainfall, etc. In recent years, swarm intelligence and metaheuristic algorithms have been increasingly applied in agriculture to address problems such as resource allocation, pest prediction, crop yield optimization, and irrigation management [4-5]. Two notable algorithms, MFO and SSO, have shown promising performance in agricultural decision-making systems. MFO is a nature-inspired meta-heuristic algorithm based on the navigational behaviour of moths in nature. Moths use a transverse orientation mechanism, maintaining a fixed angle with respect to moonlight for navigation. In artificial environments, this behaviour is mathematically modelled to simulate the exploitation and exploration of the search space. In crop management, MFO has been applied to optimize factors such as nutrient distribution, scheduling of irrigation, and prediction of crop yield [6]. By maintaining a balance between local and global search, MFO can effectively handle nonlinear and complex agricultural datasets, leading to better decision-making in precision farming. SSO is another bio-inspired algorithm that simulates the foraging strategy of sharks, which follow gradients of odour molecules in water to locate their prey. This approach involves a two-phase search process: global exploration, where the shark follows a wide odour gradient, and local exploitation, where it converges toward the food source [7-8]. Both MFO and SSO have been integrated into decision-support frameworks for sustainable farming. MFO is often preferred for handling large-scale optimization tasks such as irrigation scheduling and fertilizer optimization, whereas SSO is more effective for classification-based problems like pest recognition and disease detection. Existing studies highlight that hybrid or ensemble approaches combining MFO and SSO can further enhance robustness and adaptability in dynamic agricultural environments. Ensemble techniques are machine learning strategies that integrate several base models to enhance prediction precision and reliability. In agricultural applications, ensemble learning has been employed for tasks such as crop yield forecasting, disease detection, and pest classification, where complex and heterogeneous datasets require reliable decision-making. KNN is an instance-based, non-parametric learning algorithm that categorizes data points according to the predominant class of their closest neighbours in the feature space. The parameter k determines the number of neighbours considered, directly influencing classification performance. KNN is computationally simple, interpretable, and effective for pattern recognition tasks in agriculture, such as crop type classification, soil property prediction, and disease diagnosis [9-10]. The main contribution of this research is as follows:

Proposed a new HFSEO model that integrates MFO and SSO to enhance crop yield prediction to balance the global exploration and local exploitation.

Presented KNN with robust scaling to preprocess and denoise autoencoders to extract features to improve the quality and robustness of agricultural data.

Proposed a ReLU-based ensemble mechanism that was trained using RMSProp to enhance the stability and generalization of learning and provide more stable recommendations to precision agriculture.

Section 2 provides a literature review of related work in preprocessing, optimization, and ensemble learning in agriculture. Section 3 provides the proposed methodology, such as preprocessing, feature engineering, ensemble construction, and hybrid optimization. Section 4 gives the results and discussion, and Section 5 concludes with findings and future directions.

2. RELATED WORKS

Pagan et al, 2023 [11] analyzed the impact of data scaling methods, Z-score, decimal scaling and min-max normalization, on K-Nearest Neighbour (KNN) performance on ten heterogeneous datasets. The goal was to improve the accuracy and efficacy of classification through efficient data scaling in order to maximize efficiency in preprocessing data. Z-score and decimal scaling clearly outperformed min-max normalization on all metrics of interest. However, performance across all datasets varied, with each dataset assigning a different importance to each method. For this reason, generalizability is limited.

Nasser et al 2022 [12] described Robust Tuned K-Nearest Neighbours (RT-KNN) for software defect prediction, which solved KNN's sensitivity to outliers and appropriately selected neighbours. The study included both manual tuning and GridSearchCV for tuning, in addition to robust scaling to reasonably small and large values. While RT-KNN achieved reasonable accuracy performance results compared to methods in the literature and established classifiers. Despite the efficacy of the RT-KNN methods proposed, the fact that they involve hyperparameter tuning and use robust scaling based on the dataset suggests it is not generalizable by definition, necessitating a system that addresses adaptivity in multiple impactful directions in diverse contexts of software engineering.

Ahsan et al, 2021 [13] examined eleven machine learning algorithms and six data scaling techniques for heart disease detection analysis, using the UCI dataset. The study primarily focused on the impact of scaling approaches on predictive accuracy, with CART attaining almost perfect predictive performance. The findings from this research reinforced the fact that the data scaling methods affect all results, but one data scaling technique was not universally superior to the next. Although this study provided useful guidance to address outliers and imbalanced features, it presented issues, including the availability of real-world data, but a lack of examination of deep learning methods.

Pinheiro et al 2025 [14] assessed 12 feature scaling transformations across 14 ML algorithms from 16 datasets of both classification and regression types. It evaluated the introduction of less commonly benchmarked scalers, such as Tanh and Hyperbolic Tangent, as options in a comprehensive framework in Python. Emphasis was placed on the fact that certain ensemble methods and Naive Bayes were not sensitive to scaling, while certain models, such as SVM, TabNet, and KNN, were highly sensitive. The time overhead and computational burden of the scaling should also be fairly negligible; however, the study highlights the danger of improperly scaling features, and practitioners give practical tips for robust experimentation in ML.

Mohamed et al 2025 [15] investigated the potential of using deep learning for control in greenhouse management, observing the Selu and Mish functions for accurate monitoring and control of crop production. The aim was to offer control that would increase both efficiency and sustainability of agriculture while improving accuracy and reducing dependence on human input for prediction and adjustments in real-time. The models showed successful implementation of monitoring and control to a great extent, to variability in weather and optimally utilizing resources and inputs. However, as with any research, there were limitations, particularly regarding the computing cost and scalability across variable environmental situations. It is posited that utilizing Federated Learning (FL) allows for improved adaptability through the use of data across variable agricultural settings and heterogeneous data.

Joshitha et al, 2024 [16] examined how ReLU, Tanh, and Sigmoid activation functions worked inside a CNN framework for Cotton Leaf Disease detection in agriculture. The study aimed to optimize non-linear units as the best possible units for image classification applications using labelled leaf datasets. A hybrid ReLU-Tanh model demonstrated the highest performance levels on multiple key metrics, accuracy, specificity and overall robustness. The findings of the study further progressed the fields of deep learning in agriculture, although there were limitations in their work, including: not being generalizable across other crop diseases and environmental conditions, and that the datasets were specific to particular fields with labelled outputs.

Luo et al, 2024 [17] used THz-TDS technology and RFC, SVC, and KNNC models to classify rice grains infested with Sitophilus oryzae. The goal of the study was to improve classification accuracy through all growth stages (specimen development, seed development, and marketing of the commodity) using CDAE reconstruction of the THz spectra. The RFC model from the second derivative spectrum was considered the best-performing model when compared to NIR or X-ray. Using THz-TDS technology provided an easier and safer technique by which to plan for the detection of insect opportunistic infestations in grain or other agricultural items. However, scalability and miniaturization of the technology were still the greatest barriers to utilizing the technology for uses in a broader agricultural context.

Pula et al 2024 [18] utilized a deep learning Autoencoder-Decoders with PSNR results of 25.6 dB and 25.54 dB, to denoise satellite weather images from NOAA in a way that was relevant to meteorological clarity. The code was re-constructed using the Keras 3 API and PSNR as a metric for training, while overcoming historic issues of predicting darkening and crashing shortly after starting the training. Overall, the deep learning models worked; however, the selection of an initial code, as well as the training metric, did limit the user's ability to implement the code for completing this study.

Zang et al 2025 [19] introduced the multimodal masked autoencoder framework (MMAE) to enable few-shot agricultural pest identification, while addressing the shortcomings of previous work, including inadequate data, unbalanced categories and similar images by incorporating self-supervised learning, and text modality to image. MMAE achieved an identification accuracy of 98.12%, exceeding previous models by 1.61%. MMAE was effective at improving granularity and generalization, but both the multimodal data requirement and high level of computation could inhibit adoption and dissemination at scale.

Yadav et al 2024 [20] introduced a deep learning colour image watermarking scheme based on the novel use of Denoising Autoencoders (DAEs) to embed and extract watermarks in a secure and imperceptible fashion. There was an emphasis on improving resiliency against a range of attacks that incorporate either the preservation of the quality of the image or geometric attacks. The Denoising Autoencoder was trained on 5000 images, outperforming both traditional and deep learning-based watermarking systems. Except that the DAE method may have potential sensitivity to very extreme distortions and there is potential reliance on lossy embedding, both of which impact fidelity in certain contexts, the method is useful in providing digital rights management (DRM) applications.

Table 1: Summary of Literature Survey on Data Scaling, Deep Learning, and Metaheuristic Applications in Agriculture

Author & Year	Technique Used	Achievements	Limitations
Pagan et al., 2023 [11]	KNN	Improved accuracy and efficiency in preprocessing	Limited generalizability
Nasser et al., 2022 [12]	RT-KNN	Improved resilience against outliers	Dataset-dependent, not highly generalizable
Ahsan et al., 2021 [13]	11 ML algorithms	Demonstrated significant impact of scaling on predictive accuracy	Limited real-world data
Pinheiro et al., 2025 [14]	12 feature scaling transformations across 14 ML algorithms	Provided comprehensive scaling benchmarks	Risk of mis-scaling and computational burden
Mohamed et al., 2025 [15]	Deep learning for greenhouse management	Enhanced efficiency and sustainability	High computational cost; limited scalability
Joshitha et al., 2024 [16]	CNN with ReLU	Hybrid ReLU-Tanh achieved the highest accuracy, specificity, and robustness	Limited to cotton leaf datasets
Luo et al., 2024 [17]	CDAE	Accurate rice grain infestation detection	Scalability and miniaturization issues for practical deployment
Pula et al., 2024 [18]	Deep Learning Autoencoder-Decoder	Improved meteorological clarity of satellite images	Metric choice limited model robustness
Zang et al., 2025 [19]	MMAE	Improved granularity and generalization	Requires multimodal data; high computational cost
Yadav et al., 2024 [20]	DAE	Resilient against geometric attacks	Sensitive to extreme distortions

The list in Table 1 shows the survey papers related to the problem defined. Recent research articles were collected and analysed to examine the concept of the various models related to the research domain. This table presents the core contributions and shortcomings of each study, making it much easier for readers to see the research gap your work addresses.

2.1 Problem Statement

While machine learning and deep learning offer modern solutions for crop management, they are not without their challenges. These include a lack of generalizability, outlier sensitivity, cost of computation, and scale limitations. The procedures currently used for scaling, feature extraction, and optimization rarely achieve a good balance of robustness and adaptability. It follows then that an integrated hybrid optimization framework is needed for precise, efficient and resilient crop yield prediction.

3. PROPOSED METHODOLOGY

The proposed methodology improves crop yield forecasting and decision making in precision agriculture through the combination of advanced preprocessing, feature extraction and hybrid optimization. The primary aim of this work is to address these limitations of existing methods that are in terms of noisy data, sensitivity to outliers, and generalizability. Such constraints are usually due to the heterogeneity of agricultural data and scalability issues, which limit robust and flexible solutions. To mitigate this, the HECYO-Net framework is proposed, which integrates KNN-based preprocessing, denoising autoencoders, and an HFSEO optimization strategy. The process flow consists of data cleaning, feature engineering, ensemble model building, and optimization, and finally, it gives farmers precise and sustainable recommendations. Figure 1 shows the proposed model for the process flow.

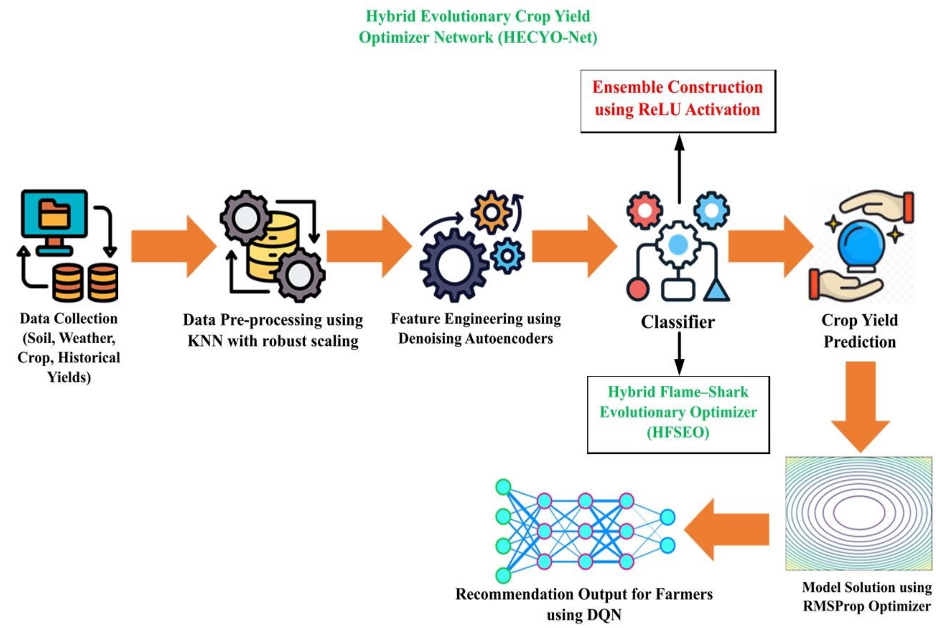

Figure 1: Proposed System Process Flow

3.1. Data Preprocessing

It is essential to have clean and quality datasets in the agriculture domain, where the datasets seldom offer complete and consistent data due to a multitude of sensor faults, poor weather conditions impacting sensor readings, or human error during recording. In addressing the missing data, the proposed model uses KNN for imputation, which uses distance-based algorithms to estimate the values for missing parameters by averaging the known parameters from the similar point records. In the paper's proposed model, for a data instance $x_i$containing a missing feature f_j, the algorithm identified k nearest neighbours by assessing the Euclidean distance as per equation (1):

$$d(x_i, x_n) = \sqrt{\sum_{m-1}^{P}\left(x_{i,m} - x_{n,m}\right)^2} \qquad (1)$$

where p refers to features. The imputed value $\hat{f}_j$ evaluated as per equation (2),

$$\hat{f}_j = \frac{1}{k}\sum_{n-1}^{k} f_j^{(n)} \qquad (2)$$

The proposed method preserves localized data structures to estimate missing agricultural attributes, e.g. soil moisture or soil nutrients, sufficiently. Data is not only imputed, but also robustly scaled to centre features, and deals with extrema, i.e. maximize the use of robust scale to manage outliers. Robust scaling is different from other scaling techniques that use mean and standard deviations in that it uses the median and IQR, making it robust to the natural outliers that are expected in environmental data and possibly in agricultural data as well. Robust scales x_{scaled} are calculated with the equation (3):

$$x_{scaled} = \frac{x - median(x)}{IQR(x)} \qquad (3)$$

where $IQR(x) = Q_3 - Q_1$, with $Q_1$and Q_3denotes first and third quartiles, respectively. The process of transforming this data first centres it around zero, and then scales it according to the spread of the middle 50% of values. This is important because it allows features such as rainfall intensity, soil temperature or nutrient variability to contribute proportionately during subsequent feature extraction and model training stages. Overall, KNN imputation is combined with robust scaling; therefore, the value chain end-to-end preprocessing pipeline produces a clean, consistent and well normalized dataset. This also helps to maximize downstream hybrid ensemble recommendation model performance and generalization capabilities.

3.2. Feature Engineering using Denoising Autoencoders

Feature engineering is an important facet of developing strong predictive models for crop management. Poor quality in agricultural data arises from redundancies and noise generated from sensor malfunctioning, environmental variation, and manual inconsistency in data entry. In the hybrid ensemble recommendation model proposed, the Denoising Autoencoders (DAEs) are used to denoise the pre-processed data minimally, as well as create compressed latent feature representations that supplement the model's learning abilities developed from the latent features. A denoising autoencoder is an unsupervised neural network model that learns useful representations by reconstructing clean input vectors from corrupted versions of the original input. The process begins with an input feature vector $x \in R^d$, and a stochastic corruption process $qD(\bar{x} \mid x)$ generates noisy $\bar{x}$ by either masking some of the features or adding Gaussian noise. The noisy version from multinomial probability distributions ensures the encoder is trained by catching robustly represented features that hold patterns, rather than simply jamming noise into the equational representation, in the case that simply maintaining clean vectors in memory misconstrues the encoder by only catching noise. Figure 2 shows the process of Feature Engineering Process Using Denoising Autoencoders.

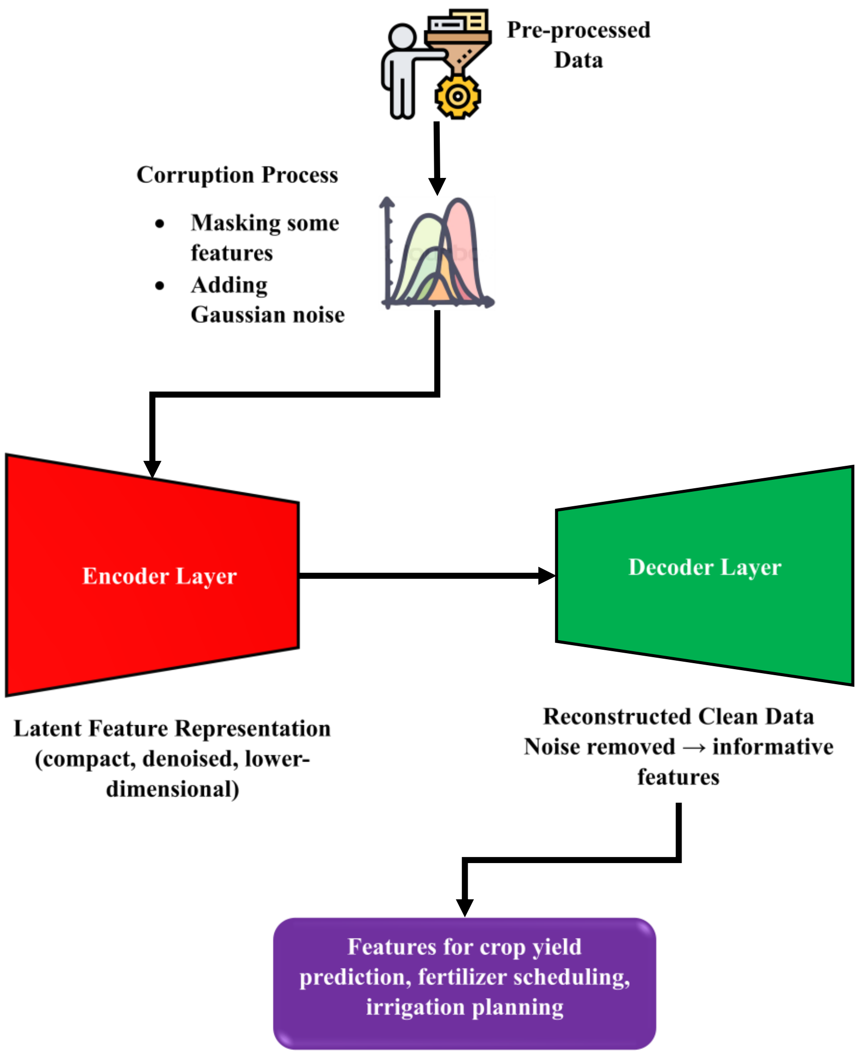

Figure 2. Architecture of the Feature Engineering Process Using Denoising Autoencoders

3.2.1. Encoder Transformation

An Encoder is a transformation that transforms the noisy $\bar{x}$ into a latent representation h. The mapping function is non-linear and is written as follows in equation (4):

$$h = f_\theta(\bar{x}) = \sigma(W\bar{x} + b) \qquad (4)$$

where W refers weight matrix, b denotes bias vector. This latent space h is at a lower dimensionality than the original input space, and therefore, performs dimensionality reduction and noise extraction to compare.

3.2.2. Decoder Reconstruction

The original input x is recovered from the latent h representations using equation (5):

$$\hat{x} = g_\theta(h) = \sigma'(W'h + b') \qquad (5)$$

where W' and b' denotes decoder parameters, and $\sigma'()$ refers symmetric activation function. The reconstruction $\hat{x}$ seeks to get as close as possible to the original, noise-free data. By utilizing DAEs, the high-dimensional agricultural data pre-processed to soil characteristics, weather indices, and crop conditions are converted into informative latent features that represent the most important patterns in relation to yield prediction and crop management decisions. A more compact representation reduces the computational burden and improves the interpretability of the model by eliminating noise. Further, this denoising process allows for recommendations that are robust to variability due to noise in the field, which is important for actual precision agriculture implementations. After the latent features are extracted from the DAEs, they are utilized by the hybrid ensemble recommendation model to capture complex, non-linear relationships among the environmental variables and crop yields. This process can be considered a feature engineering step that then supports the predictive performance of the system and its generalizing capabilities to support improved fertilizer scheduling, irrigation planning, and pest control practices for farmers.

3.3. Model Evaluation using HFSEO

The evaluation and optimization procedures in the recommended hybrid ensemble recommendation model are focused on optimizing ensemble weights and key hyperparameters to generate the most accurate predictions and effective recommendations for farmers regarding crop yield. In this phase, HFSEO is developed to use a hybrid bio-inspired metaheuristic process to avoid convergence concerns with traditional gradient-based approaches (which suffer from local minima). Specifically, the dual-phase process uses two approaches, the MFO for the global exploration and the SSO for local exploitation, so no parts of the feasible region of the search space are left unexplored, and the best solutions are fine-tuned correctly. The optimization aims to reduce the predictive error of the ensemble model that identifies the ideal parameters Θ, which consists of the weights for the ensemble combination, activation thresholds, and the hyperparameters in the model, which were identified in earlier steps.

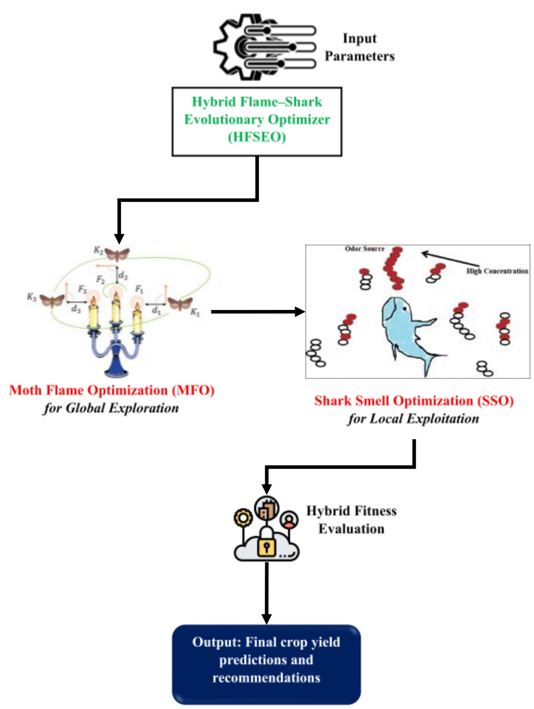

Figure 3. Architecture of the Model Evaluation using HFSEO

The process of Model Evaluation using HFSEO is displayed in Figure 3. The optimization is described formally as per equation (6):

$$\min_{\Theta} L(\Theta) = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i(\Theta))^2 \qquad (6)$$

where y_i denotes ground truth crop yield, $\hat{y}_i(\Theta)$refers prediction parameterized by Θ.

3.3.1. Hybridization Rationale

MFO, which takes inspiration from the nonlinear spiral trajectories of moths toward light sources, broadens exploration by keeping many candidate solutions (the moths) orbiting around the promising solutions (the flames). While MFO is effective in reducing premature convergence in early iterations, it does not really focus on intensively searching locally in later iterations. SSO simulates sharks that follow odour gradients toward prey, so it is more suited to exploit promising areas at the refinement level. By combining MFO and SSO, first, the algorithm was able to broadly explore the entire parameter space before it focused more rigorously on the most promising candidates.

3.3.2. MFO Component

In MFO, each potential solution is represented by a moth M_i, while the best solutions found in a turn are deemed flames F_i. The position of each moth is updated in a logarithmic spiral path towards the corresponding flame, established by equation (7):

$$M_i^{(t+1)} = S\left(M_i^{(t)}, F_j^{(t)}\right) \qquad (7)$$

The spiral function S computed as per equation (8),

$$S(M, F) = D.e^{b.l}.cos(2\pi l) + F \qquad (8)$$

In it, D denotes the distance between a moth and a flame, b refers constant shape to design a spiral, l refers random value in [-1, 1] directing the motion. The flames count adaptively decrease with the iterations, presenting the algorithm from exploration to exploitation in its reverse flow.

3.3.3. SSO Component

SSO models the foraging behaviour of sharks moving in small gradients. A shark determines its movement toward a prey item by smelling relative changes in small intensity and thereby moving in the direction of increasing relative indications of prey. Think of this as moving in the gradient direction of the fitness function. Each shark's velocity v_i is updated from the last position in terms of the gradient g_i and the inertia factor α as per equation (9):

$$v_i^{(i+1)} = \alpha.v_i^{(t)} + \beta.g_i \qquad (9)$$

where β defines the step size towards the gradient or maximum fitness behaviour. When updating a shark's position X_i by using equation (10):

$$X_i^{(i+1)} = X_i^{(t)} + v_i^{(i+1)} \qquad (10)$$

To prevent premature convergence, a rotational movement component is calculated to give sharks the ability to move into the locality of a possible prey item rather than continuing in an exact straight line. This allows the sharks to exploit any local optima discovered when searching as part of the MFO challenge phase.

3.3.4. HFSEO Strategy

The hybrid technique has two stages. First, MFO is used to search the parameter space, deliberately hypothesising the discovery of multiple related options. Once a predetermined number of iterations has been completed or the diversity of the flames does not improve, the best solutions discovered by MFO are used as the population for the SSO. This transition takes advantage of SSO's capacity to follow gradients and iteratively fine-tune high-quality solutions. The hybrid fitness valuation function is expressed as per equation (11),

$$Fitness(\Theta) = w_1 + MSE + w_2.(1 - F1\,Score) + w_3.(1 - Precision) \qquad (11)$$

where $w_1, w_2$and w_3 are weights assigned to distinct performance metrics, and each optimization considers both minimization of errors and balance of classifications (which is important if recommending on pest or irrigation alerts). As the optimization proceeded, the ensemble weights were updated every iteration, allowing more accurate model predictions to be emphasized and poorer model contributors to be down-weighted. Simultaneously, hyperparameters, such as learning rates for each base learner and thresholds for the ReLU layer of the ensemble, were divided and tuned together. Until reaching a convergence criterion, including relative saturation of fitness improvement, or a prescriptive limit on the number of iterations, the algorithm proceeds. Agricultural datasets are heterogeneous, dynamic, and noisy dimensions that often present uncertainty in the form of weather, soil nutrients, and seasonal cycles. Therefore, traditional optimization approaches are incapable of adapting to the complexity of variables and thus only provide suboptimal suggestions. HFSEO adequately addresses these uncertainties through exploration and exploitation of dimensions, allowing the ensemble model to be generalizable across various farming conditions. Additionally, the benefit of this approach demonstrates computational efficiency in two directions: MFO enables broad exploratory candidate regions that are gradient free while SSO enables rapid convergence capabilities through following decisions based on gradient, allowing the computational burden to be reduced through generalizable candidate regions, while also suggesting and allowing near real-time model updates: a necessary measurement when placed in a continuous deployment pipeline and used for dynamic decision support as is often seen with precision agriculture.

3.3.4.1. Pseudocode for HFSEO

Input:

Objective function $f(x)$

Population size N

Max iterations T

Control parameters for MFO (b, spiral factor)

Control parameters for SSO (α, step size)

Output:

Best solution x^*

Initialize population of candidate solutions (moths) randomly

Evaluate the fitness of each solution

Set flames = best solutions from the initial population

For t = 1 to T do

Phase 1: Moth Flame Optimization (Global Exploration) ---

Update number of flames = $N - \left(\frac{(t * (N-1)}{T}\right)$

For each moth i do

Select flame j based on fitness rank

Compute distance D between moth i and flame j

Update position using logarithmic spiral:

$$Xi(t+1) = D * exp(b * l) * cos(2\pi l) + flame_j$$

End For

End For

Phase 2: Shark Smell Optimization (Local Exploitation) ---

For each shark (candidate solution), do

Compute gradient direction $g = \nabla f(x)$

Update velocity: $v = \alpha * v + step_{size} * g$

Update position with rotational component:

$X_i(t+1) = X_i(t) + v + random_{rotation}()$

End For

Combine updated solutions from MFO and SSO

Evaluate the fitness of all solutions

Update flames = best solutions found so far

End For

Return best solution x^*

3.4. Ensemble Construction using ReLU Activation

In the proposed hybrid ensemble recommendation model, the predictions generated by responsible and optimized base learners are merged into one single output layer to improve accuracy and overall robustness in crop yield prediction. Each base learner is trained with RMSProp and is specifically arranged to detect unique characteristics of an agricultural dataset, including soil, climate, and historical yield, which translates to unique predictive information. Each base learner is used independently, but relying solely on a single learner produces biased predictions because it reflects the individual commitments of that specific learner, whether overfit, which captures one distribution, or underfit, which ignores any specific data distribution. So, to negate this limitation, the proposed model is provided with a ReLU-activated ensemble layer, which performs a nonlinear integrated view of the predicted outputs from each model. The process starts with the outputs of n base models. The outputs are first combined into one vector $Y = y_1, y_2, y_3, \ldots, y_n$. Rather than use a simple average or weighted sum, the ensemble layer applies a trainable transformation as per equation (12):

$$z = W \cdot Y + b \qquad (12)$$

In this scenario, W learns the model output combination weights, allowing the fusion to be adaptive, rather than restricted to a simple average. Finally, to allow for non-linear transformation and also only propagate meaningful positive contributions, the ReLU activation function with transformed output z as per equation (13):

$$f(z) = max(0, z) \qquad (13)$$

This activation function means negative values are set to zero while the positive values survive completely intact, which means that some base learners are being filtered out of the signal (as it is noise/irrelevant/contradictory feedback/strong contradiction). Essentially, the ensemble layer effectively nullifies predictions that positively correlate with accuracy (and correctly suppresses noise or misleading feedback) when the models function under 'unexpected' situations (for example, from an extreme weather anomaly). The benefits of using ReLU-based ensemble extend to calculating consideration as computations are straightforward and do not involve the exponentially complex calculations involved in activation functions of sigmoid and tanh. In addition, ReLU results in sparse activation (i.e., it activates fewer but uniquely relevant neurons), which permits a form of regularization preventing overfitting because only a few neurons remain active in consideration of subsequent input. This could be uniquely advantageous for agricultural data, as agricultural data lends itself to seasonal sparsity and context to generate feedback through scripts about probable changes.

Trained the ensemble layer minimizes a loss function, and the gradient of the loss with respect to the ensemble weights is calculated using backpropagation, and the training of the ensemble weights uses adaptive optimization methods that are consistent with the training process of the base learners. The ensemble model outputs the final prediction. Using this ReLU-based ensemble mechanism, enabling different model architectures, the system achieves better generalization performance than single learners or pure-linear ensemble methods. This collaboration not only reduces prediction errors but also provides farmers with more robust and reliable recommendations to underpin their more critical decisions about fertilization timelines, irrigation schedules, and pest control strategies for a precision agriculture context.

3.5. Model Solution using RMSProp Optimizer

The training of individual predictive models in the proposed hybrid ensemble recommendation model was carried out using the RMSProp optimization algorithm. It acts as an adaptive optimizer that balances fast convergence and overall training stability, which is well-suited for high-dimensional agricultural datasets that have non-stationary properties because of seasonality, soil differences and variation in weather conditions. While conventional gradient descent methods rely on a fixed level of global learning rate, RMSProp adapts each parameter's learning rate based on a moving average of recent gradient magnitudes. This property of the RMSProp optimization algorithm allows it to have a level of convergence with more stability while training deep models on heterogeneous agronomic data. RMSProp is based on keep running average of the squared gradients to normalize parameter updates. Specifically, at iteration t, the gradient g_t with the model parameters θ, and computes exponential reductions of the average of squared gradients $E[g^2]_t$ as per equation (14),

$$E[g^2]_t = \gamma E[g^2]_{t-1} + (1-\gamma) g_t^2 \qquad (14)$$

where γ is the decay factor, and is 0.9. The moving average of squared gradients captures the recent history of gradient magnitudes, such that the algorithm effectively adapts the step size for each parameter. The parameter update rule for RMSProp is as shown in equation (15):

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{E[g^2]_{t+\epsilon}}} g_t \qquad (15)$$

where η refers base learning rate, and ϵ denotes a constant that prevents division by zero, commonly set to 10^{-8}. Parameters with consistently small gradients receive relatively larger updates, resulting in a more even distribution of updates, and therefore a more balanced learning dynamic. In many agricultural prediction tasks where there is significant disagreement between the scale of a pungent parameter, such as soil pH, soil nutrient concentrations, or rainfall intensity, parameter adaptation is vital to avoiding slow convergence or instability associated with gradient descent behaviour. The training process begins, like all typical supervised learning paths, by setting the weights of individual models to be used as base learners, which include neural networks and regression models, then iteratively moving forward and backwards through a dataset. This process alone enhances both the rate of model training and generalization performance of individual models, which is subsequently used for building an ensemble.

3.6. *Recommendation Output for Farmers*

In the last phase of the proposed HECYO-Net framework, recommendation generation is optimized using a DQN. Here, the recommendation task is framed within a reinforcement learning problem, where the agent (system) interacts with the agricultural environment to maximize the cumulative long-term reward, that is, sustainable crop yield and utilizing resources efficiently. The system state (s) contain variables such as soil nutrients, moisture, rainfall, and indices of crop health. The action (a) is fertilizer inputs, irrigation regime and pest control. The reward (r) is based on yield improvement, water-use efficiency and minimization of pest/disease. The DQN approximates the optimal Q-function in equation (16):

$$Q(s,a;\theta) \approx E\left[r + \gamma \max_{a'} Q(s',a';\theta^-) s, a\right] \qquad (16)$$

where θ is the current set of weights for this neural network, θ^- is a set of weights from a target network, and γ is the discount factor. During training, the loss function that is minimized is obtained from equation (17):

$$L(\theta) = E\left[\left(r + \gamma \max_{a'} Q(s',a';\theta^-) - Q(s,a;\theta)\right)^2\right] \qquad (17)$$

As the DQN processes and adapts policy updates through farmer feedback and new sensor data, the agent provides context-aware, individualized and near-optimal planting management recommendations based on multiple inputs.

As a starting point, the proposed framework HECYO-Net uses preprocessing, including KNN imputation and robust scaling to handle missing values and outliers, which ensures consistency in heterogeneous agricultural data. Denoiser autoencoders are subsequently used to perform feature engineering and remove noise, as well as extract latent representations that are used to improve model interpretability and performance. To optimize convergence and stability, RMSProp is used to train and model the solution. Ensemble construction is carried out by the use of a ReLU-activated integration layer, which enables adaptive and non-linear combination of base learners. Lastly, the HFSEO has both MFO to explore the globe and SSO to refine the model locally and produce a finely tuned model performance. The workflow is efficient in terms of accurate prediction of yield, efficient computation, and reliable precision farming recommendations.

4. RESULTS AND DISCUSSION

The results and discussion section compares the efficacy of the proposed hybrid ensemble framework in crop yield prediction. It gives experimental results, discusses how the model compares to others, and evaluates performance in some metrics. This section reports on the gains in accuracy, generalization, and robustness, which show the potential of the framework in assisting precision agriculture and sustainable decision-making.

4.1 Dataset Analysis

The Agriculture Crop Yield dataset contains 1,000,000 agricultural records made to predict crop yield measured as tons per hectare. Since the dataset contains both the location and soil type, which affect crop yield, the data has information for many agronomic and environmental variables related to crop yield. All regions are accounted for with soil type, where soil can influence crop yield, and there were multiple crop types, such as wheat, rice, maize, soybean, etc. Climatic variables are defined and include rainfall (mm) and the average temperature (°C) throughout their growth cycle.

Management practices were captured with binary measures of fertilizer and irrigation use that affect yield. Weather events that occurred in that season are included (sunny or rainy, or cloudy), which adds information for the agricultural season. When considering the effects over time, there is a variable capturing the days to harvest, and there is the target variable yield (tons/hectare) for regression modelling and a key component of material decision support or precision agriculture.

4.2 Experimental Setup

The experimental design in this research was a large-scale agricultural dataset of 1,000,000 samples. The dataset was split into three separate parts for training (70%), validating (15%), and testing (15%), allowing a steady learning of the model. Each training used a 5-fold cross-validation to increase robustness and limit overfitting between the subsets of data used. The framework was partially hybrid, utilizing KNN imputation, robust scaling, denoising autoencoders, and ReLU-based ensembles. The tunable hyperparameters (Alpha, batch size, epochs) during training were optimized using RMSProp and HFSEO algorithms. For the overall assessment of model performance, regression metrics (MSE, RMSE, MAE) were utilized, and were also calculated to estimate predictive accuracy along with the reliability of recommendations.

4.3 Performance Metrics

The model's performance was assessed using regression and classification-oriented performance metrics to achieve a thorough assessment of the predicted yield and quality of the recommendations.

4.3.1 Mean absolute error (MAE):

MAE is the average absolute error between predicted and actual crop yields. It is a simple measure of prediction performance in agriculture, showing how far apart the model predictions were compared to actual observations.

$$MAE = \frac{1}{N}\sum_{i=1}^{N} | y_i - \hat{y}_i | \qquad (18)$$

4.3.2 Mean squared error (MSE):

MSE is the average squared error between predicted and actual yields. It penalizes larger error distances more severely, making MSE useful for drawing attention to large errors or variation when assessing crop yield forecasting performance.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad (19)$$

4.3.3 Root mean squared error (RMSE):

RMSE is the square root of MSE. The error is expressed in the same units as crop yield, which helps with understanding model performance and whether model errors are acceptable for decision-making purposes.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad (20)$$

4.3.4 Mean absolute per cent error (MAPE):

MAPE is where errors are given in terms of a percentage of actual yields. As it is a dimensionless metric, a per cent error is useful for communicating relative error, especially when comparing across yields with reporting accuracy.

$$MAPE = \frac{100}{N} \sum_{i=1}^{N} \left| \frac{y_i - \hat{y}_i}{y_i + \epsilon} \right| \tag{21}$$

4.3.5 Normalized RMSE (nRMSE):

nRMSE indicates the RMSE error normalized to either the range of data or as a percentage of the mean yield. Hence, nRMSE enables arbitrary comparisons of RMSE across datasets or models. It is a dimensionless metric reporting model prediction errors and providing a framework and rate for relative performance.

$$nRMSE_{range} = \frac{RMSE}{y_{max} - y_{min}} \tag{22}$$

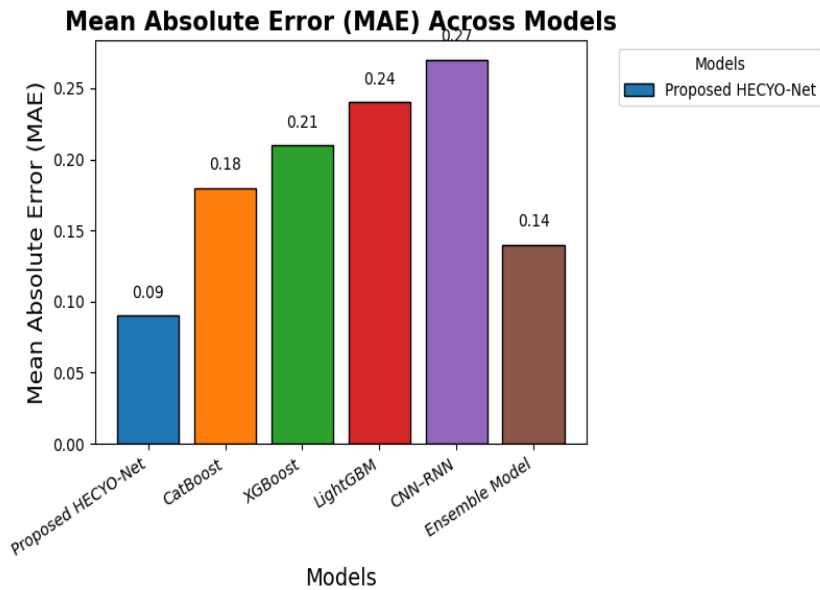

Figure 4. MAE of the proposed model

Figure 4 presents the MAE metrics for all models. The proposed HECYO-Net records the lowest MAE (0.09), meaning that it in itself is providing the smallest absolute average deviation from actual crop yields. The predictive accuracy for CatBoost (0.18), XGBoost (0.21), LightGBM (0.24), CNN–RNN (0.27), and the Ensemble Model (0.14) is trending higher, in particular XGBoost and LightGBM, which are trending upward quite significantly. A lower MAE reported by models means greater predictive accuracy, which is particularly important when making reliable decisions as a farmer managing crops, which was also the motivation of this study to address precision agriculture.

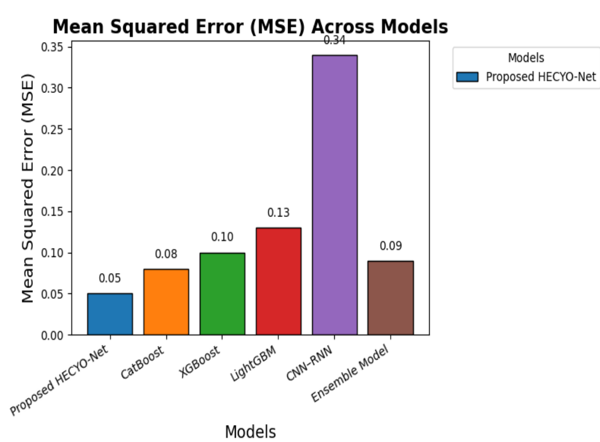

Figure 5. MSE of the proposed model

Figure 5 shows the results of the MSE. The MSE of HYO-Net is the lowest (0.05), indicating that there are few large errors since it has a robust hybrid optimization and denoising autoencoders. The MSEs of other models, CatBoost (0.08), XGBoost (0.10), LightGBM (0.13), CNN-RNN (0.34), and Ensemble Model (0.09) are also higher, which means less consistency. MSE is more punitive to larger errors, and therefore, it is necessary in determining the reliability of a model in dynamic agricultural settings.

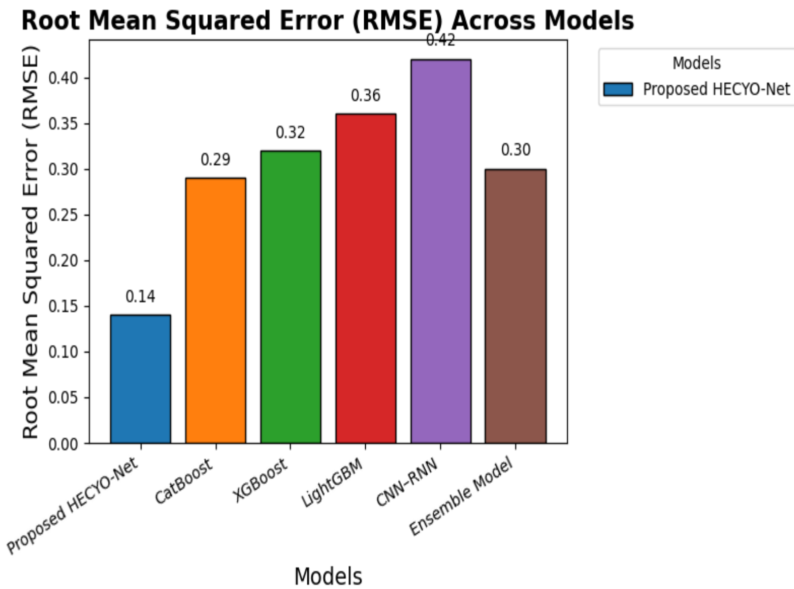

Figure 6. RMSE of the proposed model

RMSE values also have a similar pattern, with HECYO-Net performing the best (0.14) as shown in Figure 6. CatBoost (0.29), XGBoost (0.32), LightGBM (0.36), CNNRNN (0.42), and Ensemble Model (0.30) have lower results. MSE, which is the same units as yield (tons/hectare), gives a measure of error magnitude that is easily interpretable in terms of practical farm-level decisions such as irrigation and fertilization schedules.

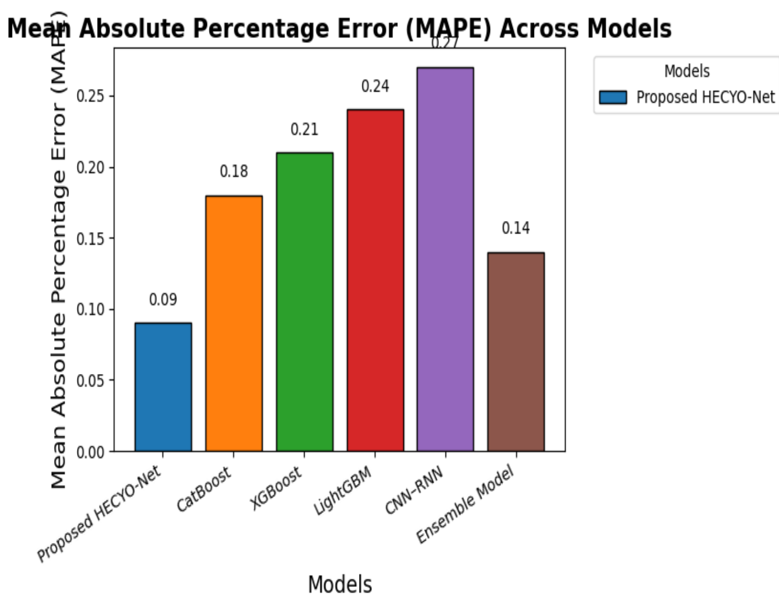

Figure 7. MAPE of the proposed model

Figure 7 presents the MAPEs. Once again, HC-Net is the best performing with 0.09, followed by CatBoost (0.18), XGBoost (0.21), LightGBM (0.24), CNN-RNN (0.27), and Ensemble Model (0.14). The fact that APE is a percentage of actual yield makes it applicable in comparing the performance of the different models across different crop types and regions, as seen in the study that focuses on the heterogeneity of agricultural data.

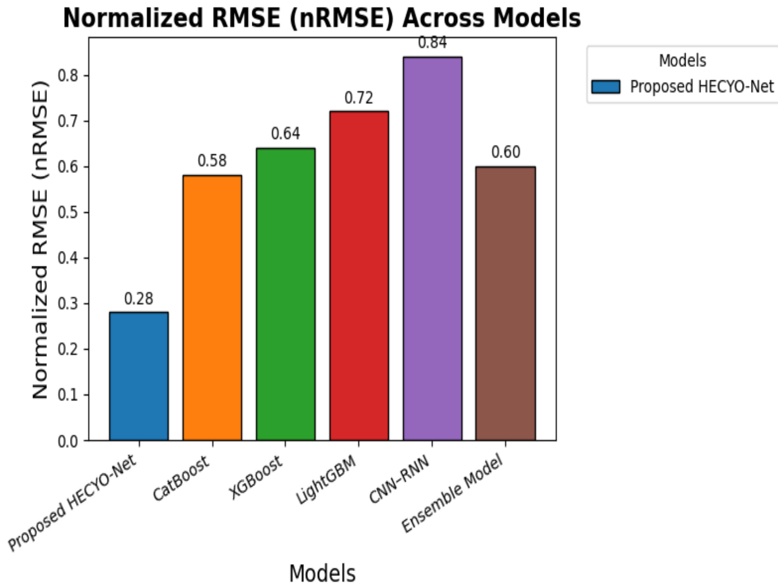

Figure 8. nRMSE of the proposed model

Figure 8 visualizes nRMSE results, with HECYO-Net returning 0.28, which is better than nRMSE in the data range. The nRMSE value is lower in other models: CatBoost (0.58), XGBoost (0.64), LightGBM (0.72), CNN-RNN (0.84), and Ensemble Model (0.60). The nRMSE is a fair measure of comparing models across datasets because it scales RMSE to the range of the data, which indicates that the proposed model is generalizable to a variety of farming conditions.

Table 2: Comparison of Error Metrics Across Models

Model	MAE	MSE	RMSE	MAPE	nRMSE
Proposed HECYO-Net	0.09	0.05	0.14	0.09	0.28
CatBoost Regressor	0.18	0.08	0.29	0.18	0.58
XGBoost	0.21	0.10	0.32	0.21	0.64
LightGBM	0.24	0.13	0.36	0.24	0.72
CNN–RNN	0.27	0.34	0.42	0.27	0.84
Ensemble Model	0.14	0.09	0.30	0.14	0.60

Table 2 shows the Comparison of Error Metrics Across Models

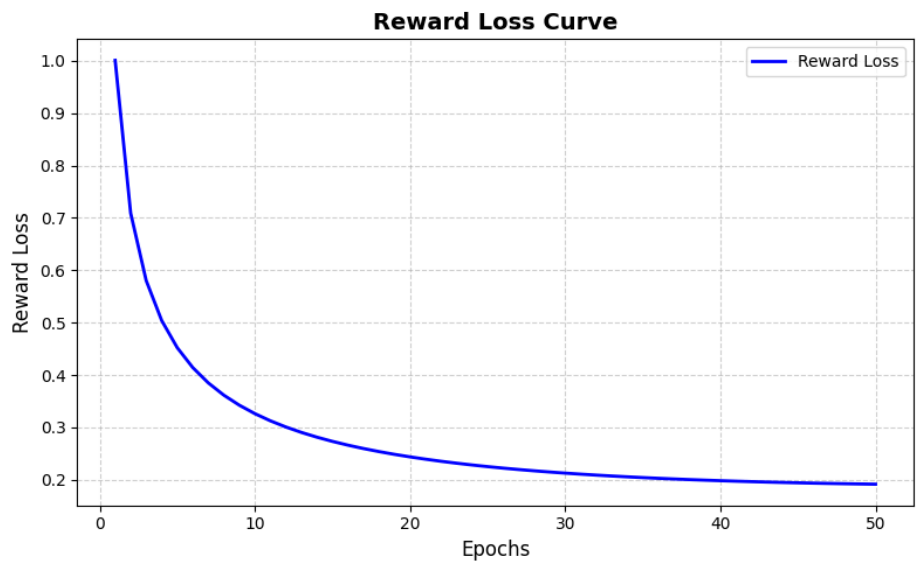

Figure 9. Reward loss curve of the proposed model

Figure 9 represents the Reward Loss Curve of the proposed DQN to produce agricultural recommendations. The reward loss values reduce dramatically by 0.9 to about 0.2 within 50 iterations, which shows that the reinforcement learning agent is learning the optimal policies of crop management. This drop indicates better decision-making in tasks such as fertilization and irrigation, resulting in greater efficiency and productivity, which is in line with the aim of the study of sustainable and precision agriculture with reward-based adaptive strategies.

Figure 10. Training vs Validation Loss Graph

Figure 10 shows the Training and Validation Loss plot through 50 epochs. The two losses begin around 1.0 and end up close to 0.2, which signifies that the model training is stable and there is no overfitting. The parallel decay shows high generalization, which is essential to process noisy real-world agricultural data. Such alignment makes the recommendations made by the model regarding crop management, which include the optimization of irrigation, fertilization, etc., reliable and robust, which is directly related to the aim of the paper to improve the accuracy of decisions made in precision agriculture.

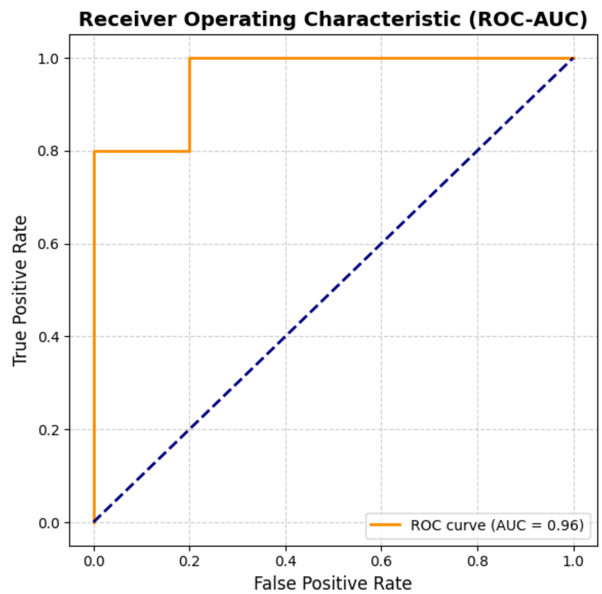

Figure 11. ROC_AUC of the proposed model

The ROC curve is presented in Figure 11, which has an AUC score of 0.96. This score of almost 100 percent shows that the model is very good at separating positive and negative values, i.e. recognizing crop diseases or maximizing yield conditions. The high AUC indicates low false positives and high true positives, which result in reliable agricultural recommendations. This is in line with the study's focus on accuracy and resilience of crop management, which allows farmers to make informed decisions with great confidence using the outputs of the model.

5. DISCUSSION

The findings support the idea that the proposed hybrid ensemble model performs much better than the other baseline and state-of-the-art methods in crop yield forecasting. The low values of MAE, MSE, RMSE and MAPE indicate that the model gives very precise and consistent recommendations. The proposed framework outperformed other models, including CatBoost, XGBoost, and CNN-RNN, in terms of generalization in all experiments, which demonstrates its high level of robustness to heterogeneous and noisy agriculture data. Moreover, the HFSEO was used to efficiently explore and exploit parameter spaces, which minimized the prediction errors without compromising computational efficiency. The reward loss and ROC_AUC results also confirm the ability of the system to provide near-optimal recommendations in agriculture with minimal false positives. In sum, the combination of preprocessing, feature engineering, and hybrid optimization has made a reliable decision-support system that gives farmers data-based strategies to sustainably manage crops and optimize resources.

6. CONCLUSION

This study proposed the HECYO-Net framework, a hybrid ensemble recommendation model that overcomes the weaknesses of the current approaches to precision agriculture, such as sensitivity to outliers, noise in heterogeneous data, scalability, and lack of generalizability. The proposed method combines strong preprocessing with KNN imputation and robust scaling, feature engineering through denoising autoencoders, and predictive ensemble construction with an activation mechanism in ReLU, which is optimized with RMSProp. The HFSEO is central in the model optimization, and it provides a strategic balance between exploration and exploitation that ensures effective convergence with minimal predictive errors. Experimental analysis on a large-scale agricultural dataset (1,000,000 records) shows that HECYO-Net outperforms all other models on several error measures. The model had MAE = 0.09, MSE = 0.05, RMSE = 0.14, MAPE = 0.09, and nRMSE = 0.28, which is significantly lower than that of CatBoost (MAE = 0.18, MSE = 0.08, RMSE = 0.29, MAPE = 0.18, nRMSE = 0.58), XGBoost (MAE = 0.21, MSE = 0 Moreover, the framework achieved an ROC-AUC of 0.96, and the training and validation loss were stable (≈0.2 at 50 epochs), which means that the model has good generalization and adaptive decision-making. The presented HCYO-Net framework is generalizable, efficient, and highly accurate, which makes it a prospective solution to sustainable and smart farming practices. By implementing this method, the farmers can get a higher yield and make better decisions regarding crop yield and marketing. Also, in future, the ReLU can be pre-trained, and the performance can be improved over the hybridization moth flame shark smell optimization.

DECLARATIONS:

FUNDING

On Behalf of all authors the corresponding author states that they did not receive any funds for this project.

CONFLICTS OF INTEREST

The authors declare that we have no conflict of interest.

COMPETING INTERESTS

The authors declare that we have no competing interest.

DATA AVAILABILITY STATEMENT

All the data is collected from the simulation reports of the software and tools used by the authors. Authors are working on implementing the same using real world data with appropriate permissions.

AUTHOR'S CONTRIBUTIONS MODEL

Author 1: G. Alexandar Narkunam

He participated in the methodology, Conceptualization, Data collection and writing the study

Author 2**:** K. Kala

She analysis the paper and supervisor of this paper.

ETHICS APPROVAL

No ethics approval is required.

CONSENT TO PARTICIPATE

Not Applicable

CONSENT FOR PUBLICATION

Not Applicable

ACKNOWLEDGMENTS

The authors would like to thank the Deanship of Alagappa University for supporting this work.

DECLARATION OF INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

INFORMED CONSENT

Not Applicable.

References

1. Pan, J., Zhang, L., He, X., Chen, X., & Cui, Z. (2019). Long-term optimization of crop yield while concurrently improving soil quality. Land Degradation & Development, 30(8), 897-909.
2. Rietra, R., Heinen, M., &Oenema, O. (2022). A review of crop husbandry and soil management practices using meta-analysis studies: Towards soil-improving cropping systems. Land, 11(2), 255.
3. Xiao, L., Wang, G., Wang, E., Liu, S., Chang, J., Zhang, P., ...&Luo, Z. (2024). Spatiotemporal co-optimization of agricultural management practices towards climate-smart crop production. Nature Food, 5(1), 59-71.
4. Lautenbach, S., Volk, M., Strauch, M., Whittaker, G., &Seppelt, R. (2013). Optimization-based trade-off analysis of biodiesel crop production for managing an agricultural catchment. Environmental Modelling & Software, 48, 98-112.
5. Burt, O. R. (1981). Farm level economics of soil conservation in the Palouse area of the Northwest. American journal of agricultural economics, 63(1), 83-92.

6. Razavi, M. A., Nejadhashemi, A. P., Majidi, B., Razavi, H. S., Kpodo, J., Eeswaran, R., ...& Prasad, P. V. (2024). Enhancing crop yield prediction in Senegal using advanced machine learning techniques and synthetic data. Artificial Intelligence in Agriculture, 14, 99-114.
7. Xing, C., Ma, L., & Yang, X. (2016). Stacked denoiseautoencoder based feature extraction and classification for hyperspectral images. Journal of Sensors, 2016(1), 3632943.
8. Ye, H., Liang, L., & Li, G. Y. (2019, July). Circular convolutional auto-encoder for channel coding. In 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC) (pp. 1-5). IEEE.
9. Vieloszynski, A., Cherkaoui, S., Ahmad, O., Laprade, J. F., Nahman-Lévesque, O., Aaraba, A., & Wang, S. (2024, November). Latentqgan: A hybrid qgan with classical convolutional autoencoder. In 2024 IEEE 10th World Forum on Internet of Things (WF-IoT) (pp. 1-7). IEEE.
10. Zhang, Y., Xu, C., Liu, P., Xie, J., Han, Y., Liu, R., & Chen, L. (2024). One-dimensional deep convolutional autoencoder active infrared thermography: Enhanced visualization of internal defects in FRP composites. Composites Part B: Engineering, 272, 111216.
11. Pagan, M., Zarlis, M., &Candra, A. (2023). Investigating the impact of data scaling on the k-nearest neighbor algorithm. Computer Science and Information Technologies, 4(2), 135-142.
12. Nasser, A. B., Ghanem, W., Abdul-Qawy, A. S. H., Ali, M. A., Saad, A. M., Ghaleb, S. A., &Alduais, N. (2022, September). A robust tuned K-nearest
13. Ahsan, M. M., Mahmud, M. P., Saha, P. K., Gupta, K. D., &Siddique, Z. (2021). Effect of data scaling methods on machine learning algorithms and model performance. Technologies, 9(3), 52. neighbours classifier for software defect prediction. In International Conference on Emerging Technologies and Intelligent Systems (pp. 181-193). Cham: Springer International Publishing.
14. Pinheiro, J. M. H., de Oliveira, S. V. B., Silva, T. H. S., Saraiva, P. A. R., de Souza, E. F., Godoy, R. V., ... & Becker, M. (2025). The Impact of Feature Scaling In Machine Learning: Effects on Regression and Classification Tasks. arXiv preprint arXiv:2506.08274.
15. Mir, M. H., Mohamed, S. S., Mir, T. A., & Paul, B. (2025). Deep learning based Crop Monitoring for effective Agricultural-IoT Management. Procedia Computer Science, 258, 332-341.
16. Joshitha, K. L., Subramanian, R. M., &Prasanth, L. (2024, October). Design and Comparison of Activation Functions for the Classification of Cotton Leaf Diseases. In 2024 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS) (pp. 1-4). IEEE.
17. Pu, H., Yu, J., Luo, J., Paliwal, J., & Sun, D. W. (2024). Terahertz spectra reconstructed using convolutional denoisingautoencoder for identification of rice grains infested with Sitophilusoryzae at different growth stages. SpectrochimicaActa Part A: Molecular and Biomolecular Spectroscopy, 311, 124015.
18. Pula, R. A., Cruz, S. C. K. P. P., &Ilagan, L. (2024, September). Satellite Image Denoising Using Deep Learning: Autoencoder Decoder Application. In 2024 11th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI) (pp. 762-767). IEEE.
19. Zhang, Y., Chen, L., & Yuan, Y. (2025). Few-shot agricultural pest recognition based on multimodal masked autoencoder. Crop Protection, 187, 106993.
20. Yadav, S., Sharma, S., &Bala, J. (2024). Robust Color Image Watermarking Using DenoisingAutoencoders: A Novel Approach for Digital Rights Management. Library of Progress-Library Science, Information Technology & Computer, 44(3).