



Ensemble Machine Learning with Optimization Driven Rice Crop Yield Prediction Model

Mr. S. Thirumal¹, Dr. R. Latha²

¹Research Scholar, Department of Computer Applications, St. Peter's Institute of Higher Education and Research, Chennai.

²Professor and Head, Dept., of Computer Applications, St. Peter's Institute of Higher Education and Research, Chennai.

Abstract

India is the second largest rice producer which contributes around 20% of the world's rice production. Rice crop is indispensable for majority of the world's population, and therefore, correct prediction of rice production is paramount for development policies, trade, decision-makers, humanitarian assistance, and so on. Crop predicting, the art of forecasting agriculture production before the crop harvest take place, assist a many interested parties making superior decisions around agricultural decision-making. Accurate, reliable, and timely rice production forecasting in India is central for global health and food security problems. Statistical machine learning models and Classical mechanistic models should recognize patterns, creating the investigation on and application of these techniques time-consuming and laborious. Therefore, this study develops a Moth Flame Optimizer with Ensemble Machine Learning Driven Rice Crop Yield Prediction (MFOEML-RCYP) algorithm. The MFOEML-RCYP methodology exploits Z-score normalization to pre-process input characteristics, which ensures optimum data dissemination for succeeding analysis. To optimize predictive accuracy, an ensemble technique is introduced, integrating the strengths of various techniques such as Backpropagation Neural Network (BPNN), Wavelet Neural Network (WNN), and Light Gradient Boosting Machine (Light GBM). Moreover, parameter optimization is implemented by the Moth Flame Optimizer (MFO) for maximizing predictive outcomes and adjusting model parameters. Empirical outcomes illustrate the effectiveness of the presented MFOEML-RCYP technique in precisely predicting rice crop yields, emphasizing its prospective for enhancing crop management approaches and informing agricultural decision-making.

Keywords: Crop Yield Prediction; Machine Learning; Moth Flame Optimizer; Fitness Function; Pre-processing

1. Introduction

The fast upsurge in the world population uses force on the region of agriculture and forbidding the food safety of the globe [1]. Between cereals, rice is the major causes of food with higher nutritive value. On-time and precise forecasts of crop yield is vital for food safety and executive organization, particularly in the present constantly varying global surroundings and condition. Dissimilar techniques have been accepted for exact yield

assessment and every model contain its individual limitations and strengths [2]. For example, the traditional area analyses and crop statistics are beneficial for accurately assessing crop harvest; but, if crop yield forecast of the larger area is preferred, the surveys show insufficient owing to time, economical, and huge experienced manpower restraints [3]. Precise crop yield predicting may also allow better plan of protection products that diminish weather risks and steady farmer profits. Weather-based crop assurance, for example, utilizes a climate index like total rainfall to define costs to farmers, meaning that insurance companies do not want to visit farmers to evaluate damages and judge claims [4]. Rather, if the climate grasps a definite threshold, quick automatic expenses can be distributed to agriculturalists, who evade the essential to sell possessions to endure owing to adverse weather actions [5].

The require for precise data on crop yields is mainly vital in countries like India, where the farming sector offers employments for more than millions of farmers, with 70% of rural homes dependent on farming for their foremost cause of income [6]. One of India's main basic crops is rice, which donates to 30% of calories used up in India and is an important trade product for the country [7]. India cultivates rice on nearby 45 million hectares of land, with a complete creation of 178 million metric tonnes [8]. Furthermore, the distribution of rainy season, which is a main source of water for rice agriculture, has become changeable in current years owing to weather variability. In such cases, crop yield forecasts might be capable to increase agricultural early warning system (AEWS) that provide innovative notification of potential dangers to crop production, allowing pre-emptive action in affected areas [9]. Preceding research has exposed that present agricultural monitoring methods lack strong crop yield and crop production predictions, as well as the operationalisation of such model at measure [10].

This study develops a Moth Flame Optimizer with Ensemble ML Driven Rice Crop Yield Prediction (MFOEML-RCYP) algorithm. The MFOEML-RCYP system exploits Z-score normalization to pre-process input characteristics, which ensures optimum data dissemination for succeeding analysis. To optimize predictive accuracy, an ensemble technique is introduced, integrating the strengths of various techniques such as Backpropagation Neural Network (BPNN), Wavelet Neural Network (WNN), and Light Gradient Boosting Machine (Light GBM). Moreover, parameter optimization is implemented by the Moth Flame Optimizer (MFO) for maximizing predictive outcomes and adjusting model parameters. Empirical outcomes illustrate the effectiveness of the presented MFOEML-RCYP technique in precisely predicting rice crop yields, emphasizing its prospective for enhancing crop management approaches and informing agricultural decision-making.

2. Related Works

Hoque et al. [11] introduced a new crop yield forecast method which uses a year's value of climatological data, crop yield data, pesticide records, and ML models. This method used rigorous models and assessed 3 ML techniques such as K-Nearest Neighbors, Gradient Boosting (GB), and Multi-variate Logistic Regression. This model also used the GridSearchCV technique for hyper-parameter change to recognize the appropriate hyper-parameter during the K-Fold cross-validation. In [12], an automatic disease analysis method for maize plants was projected. Generally, 4 phases are projected like pre-processing the data of input, splitting the affected

regions, removing the features and forecast of virus. At this paper, supervised ML models were employed in order to forecast the illnesses of maize plants. YOLO structure was employed for the segmentation procedure, while Discrete Wavelet Transform (DWT) has been utilized for removing the feature. In [13], a Robust Optimized ANNs (ROANNs) technique with GA and Multi Objective PSO (MOPSO) is projected in this work. The input parameters were enhanced both by utilizing GA and MOPSO optimizer techniques in order to rebuild the database. Rebuilt database enhanced by employing ANN- Back Propagation (BP) model. The main purpose for enhancing the development of paddy was recognized utilizing the output of Neural Networks.

In [14], a robust forecast technique is projected by employing ML techniques. The developed plan contains gathering a lot of information on variables from the objective areas. Utilizing careful feature engineering and data pre-processing, the projected method discovers related forecasters and originate major insights for rice yield assessment. Then, advanced ML approaches namely SVM, RF, and GB are used, to construct reliable forecast methods. Reyana et al. [15] presented a new Multi-sensor ML Approach (MMLA) model for categorizing multi-sensor data. Depend upon the projected recommendation model, 8 crops were categorized. Crop types were categorized utilizing 3 ML techniques like J48 Decision Tree, Hoeffding Tree, and RF. Devi et al. [16] projected an advanced DL technique which accurately takes and combines spatial and temporal features. This method predictions crop yields with a least rate of error, leveraging the sturdiness of an exclusive WaveNet and LSTM hybrid structure, presenting a fresh viewpoint to farming yield forecasts. The innovation of the technique lies in its dual-tier strategy: the 1st stage contains pre-processing which is parallel to current methods, and the 2nd stage attaches the united power of LSTM and WaveNet for feature regression and extraction, allowing exact forecasts.

3. The Proposed Model

In this study, a novel MFOEML-RCYP technique is developed. The main purpose of the MFOEML-RCYP model comprises three different kinds of processes namely Z-score normalization, ensemble learning, and parameter optimizer. Fig. 1 represents the entire flow of presented of MFOEML-RCYP model.

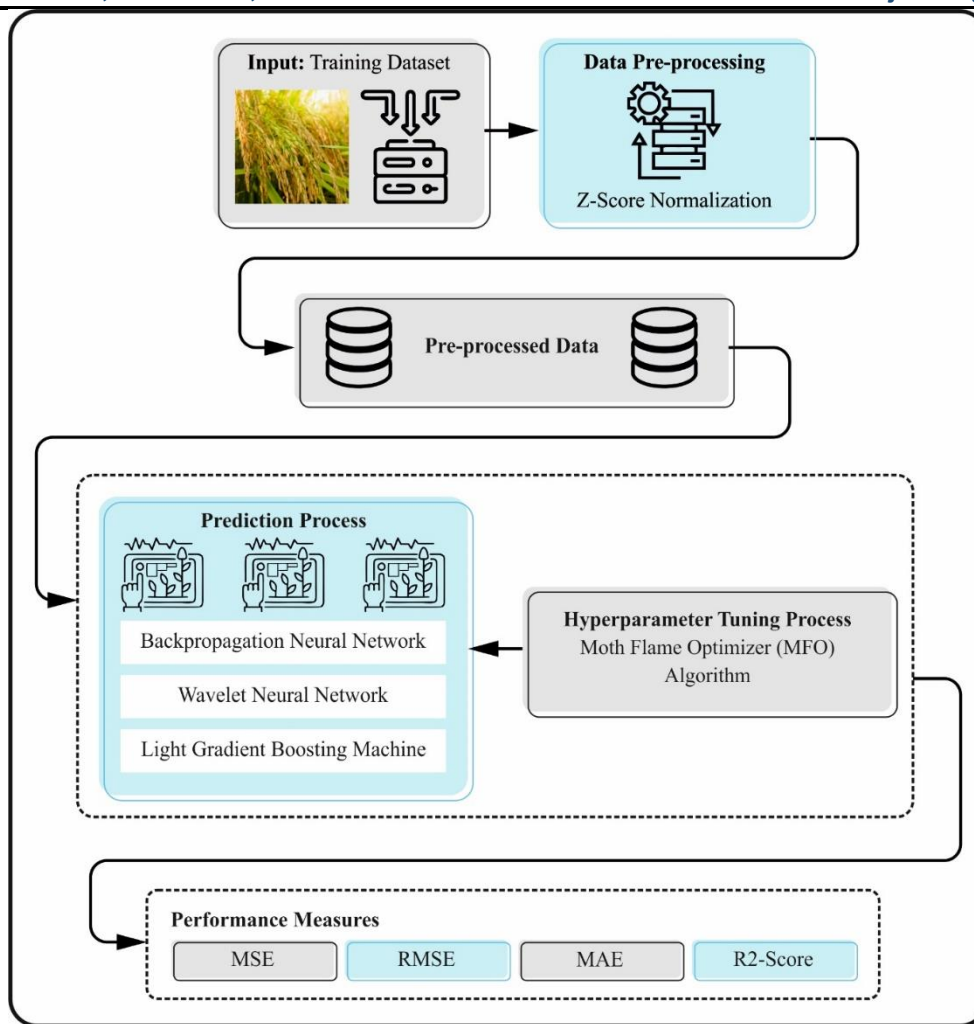


Fig. 1. Overall flow of the MFOEML-RCYP model

3.1. Z-score Normalization

Initially, the MFOEML-RCYP technique exploits Z-score normalization to pre-process input characteristics. Normalization is process of achieving scales within [0,1], whereas standardization exploits a principle named the standard deviation for describing the data distribution [17]. Computing a z-score is a standardization method, because the outcomes are outsizeing of the zero-to-one interval. Z-score normalization (standardization), perceives rescaled features in a way that follows uniform distribution properties with $\mu=0$ and $\sigma=1$, where μ refers to the mean (average) and σ denotes the standard deviation (SD) from the mean.

3.2. Ensemble Learning

To optimize predictive accuracy, an ensemble technique is introduced, integrating the strengths of various techniques such as BPNN, WNN, and Light GBM.

3.2.1. BPNN Model

BPNN is an ANN based on the error backpropagation model [18]. It is a kind of multi-layer FFNN that possesses outstanding abilities for approximating arbitrary function and non-linear mapping. Therefore, it finds wider application in pattern recognition, classification, regression, and signal processing fields. BPNN typically comprises of three layers: the hidden, the input, and the output layers.

In this study, similarity measure represented as X_1 to X_n are inputted to the network.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

The parameter χ in the sigmoid function ranges from negative to positive infinity. Z_k denotes the predictable rate according to Y_k , then the error function of BPNN is represented as follows:

$$E = \frac{1}{2} \sum_{k=1}^m (Z_k - Y_k)^2 \quad (2)$$

The BPNN adjusts W_{ij} and W_{jk} values according to the partial derivative with regard to output E for minimizing the error. Training stops once the BPNN completes the predictable amount of training iterations or the error reaches the predefined target.

3.2.2. WNN Model

The WNN is adaptably generated the wavelet templates parameters [19]. By comparison with generally used NN like BPNN, the activation function can be replaced by utilizing the collection of wavelet functions that have been produced through the Morlet wavelet generation.

The architecture of WNN comprises a 3 layers. It comprises K nodes from the input and output layers, and L nodes in the HL. The parameter in the HL that could be constantly controlled in the learning method, are consists of the weights w_{il} , w_{lo} , the wavelet translation variable b , and wavelet dilation factor a . The y_i of the output layer will be compared with input data x_i and represented as given below:

$$net_l = \frac{w_{il}x_i - b_l}{al} \quad (l = 1, 2, \dots, L) \quad (3)$$

$$x_l = \psi_{al, b_l}(net_l) \quad (l = 1, 2, \dots, L) \quad (4)$$

$$y_i = \sum_{l=1}^{l=L} w_{lo} x_l \quad (i = 1, 2, \dots, k). \quad (5)$$

$\psi(x)$ means the generative function that has been described as the Morlet wavelet function, and extensively employed as an activation function in NNs.

$$\psi(x) = \cos(1.75x)\exp(-0.5x^2). \quad (6)$$

3.2.3. Light GBM Model

It can be generated the outcomes more rapidly thereby it can be followed through the word 'Light' [20]. It will be handled the large quantity of data when employing lesser memory. The best fit splits the tree leaf-wise, while alternative boosting techniques can be divided the tree level-wise. This method reduces losses when related to the level-wise technique while developing under the similar leaf. Additional cause for its popularity have the

efficiency in the outcomes. This can be allowed the GPU learning. Hence, data scientists often applies it for developing applications in the data science field.

$$\hat{S}(b) = \frac{1}{n} \left(\frac{\left(\sum_{y_j \in q_m} h_j + \frac{1-c}{d} \sum_{j \in p_m} h_j \right)^2}{n_m^i(b)} + \frac{\left(\sum_{y_j \in q_n} h_j + \frac{1-c}{d} \sum_{j \in p_n} h_j \right)^2}{n_n^i(b)} \right) \quad (7)$$

Here $S(b)$ means the predicted difference gain in the subset $q \cup p, q_l = \{y_i \in q: y_{ji} \leq b\}, q_n = \{y_j \in q: y_{ji} > b\}, p_m = \{y_j \in p: y_{ji} \leq b\}, p_n = \{y_j \in p: y_{ji} > b\}$, and the coefficients $\frac{1-c}{d}$ will be implemented to decrease the summation of the gradients through B back to the dimension of q^c . The prediction $S(b)$ values are utilized under smaller data values rather than the accurate $S(b)$ over all events that can be employed for identifying the split point.

3.3. Parameter Optimizer

Finally, the parameter optimization is implemented by the MFO for maximizing predictive outcomes and adjusting model parameters. MFO's motivated mainly from a navigation model termed transverse positioning once moths fly during night that creates the flying path of moths at a stable viewpoint with esteem to the moon [21]. If the moon is far away, so its emissions are equivalent to the moth, which safeguards the straight flight. When a moth grabs an imitation light like bulb, it will fly near the bulb while the blub light will spread in all ways. The mathematical MFO method is defined as follows: MFO is a population-based intelligence optimizer technique that pretends the spiral match of moths near the flames. It is noticeable that flames and moths are results of the function. The dissimilarity among them is the given direction and are upgraded in every iteration.

Let OM and M matrix signify the location and fitness values (FVs) of moth individuals, separately. The F matrix signifies the flames location, and OF is the flames FVs. These matrices are conveyed as below:

$$M = \begin{bmatrix} m_{1,1} & m_{1,2} & \dots & m_{1,d} \\ \vdots & \ddots & & \vdots \\ m_{n,1} & m_{n,2} & \dots & m_{n,d} \end{bmatrix}, OM = \begin{matrix} OM_1 \\ \vdots \\ OM_n \end{matrix} \quad (8)$$

$$F = \begin{bmatrix} f_{1,1} & f_{1,2} & \dots & f_{1,d} \\ \vdots & \ddots & & \vdots \\ f_{n,1} & f_{n,2} & \dots & f_{n,d} \end{bmatrix}, OF = \begin{matrix} OF_1 \\ \vdots \\ OF_n \end{matrix} \quad (9)$$

where n, d signify the amount of moths and dimensions, correspondingly

Moth search mechanism

To pretend the tool of a moth logarithmic spiral near a fame, utilize Eq. (10) in order to upgrade the location of the moths.

$$S(M_i, F_j) = D_i \cdot e^{bt} \cdot \cos(2\pi t) + F_j \quad (10)$$

where M_i represents the i -th moth, F_j denotes the j -th fame, D_i said to be complete distance among M_i and F_j , b is a constant that disturbs the outline of the logarithmic spiral, t refers the arbitrary number from 1 and 1. When $t = 1$ signifies that the moth is extreme from the fame, and once $t = -1$ denotes that the moth is near the fame. D_i is expressed as follows:

$$D_i = |F_i - M_i| \quad (11)$$

Flame renewal mechanism

Eq. (10) is a mathematical method of each moth fly's to its corresponding fame. But, the moths upgrading their location with esteem to n dissimilar flames may decrease the exploitation possibility. To resolve this issue, an adaptive fame rule tool has been developed, which decrease the amount of flames NO_f , as presented in Eq. (12):

$$NO_f = \text{round} \left(N_{\max} - I \cdot \frac{N_{\max} - 1}{I_{\max}} \right) \quad (12)$$

Whereas, I signifies the current amount of iterations, N_{\max} denotes the maximal flame counts, and I_{\max} represents the maximal iterations count. At the time of each iteration, fames are defined and organized affording to the FVs once the fame list can be upgraded, and then it flies to the equivalent fames in order to upgrade their locations. The first moth often upgrades its location affording to the optimal fame position, while the last moth permits to the location of the worse fame. At last iteration, the moths upgrade their locations with esteem to the finest fame.

The fitness choice is a main aspect controlling the efficiency of MFO algorithm. The hyperparameter tuning method contains the encoded process to measure the solution of candidate outcomes. During this case, the MFO system assumes accuracy as a primary condition for designing fitness function (FF).

$$\text{Fitness} = \max(P) \quad (13)$$

$$P = \frac{TP}{TP + FP} \quad (14)$$

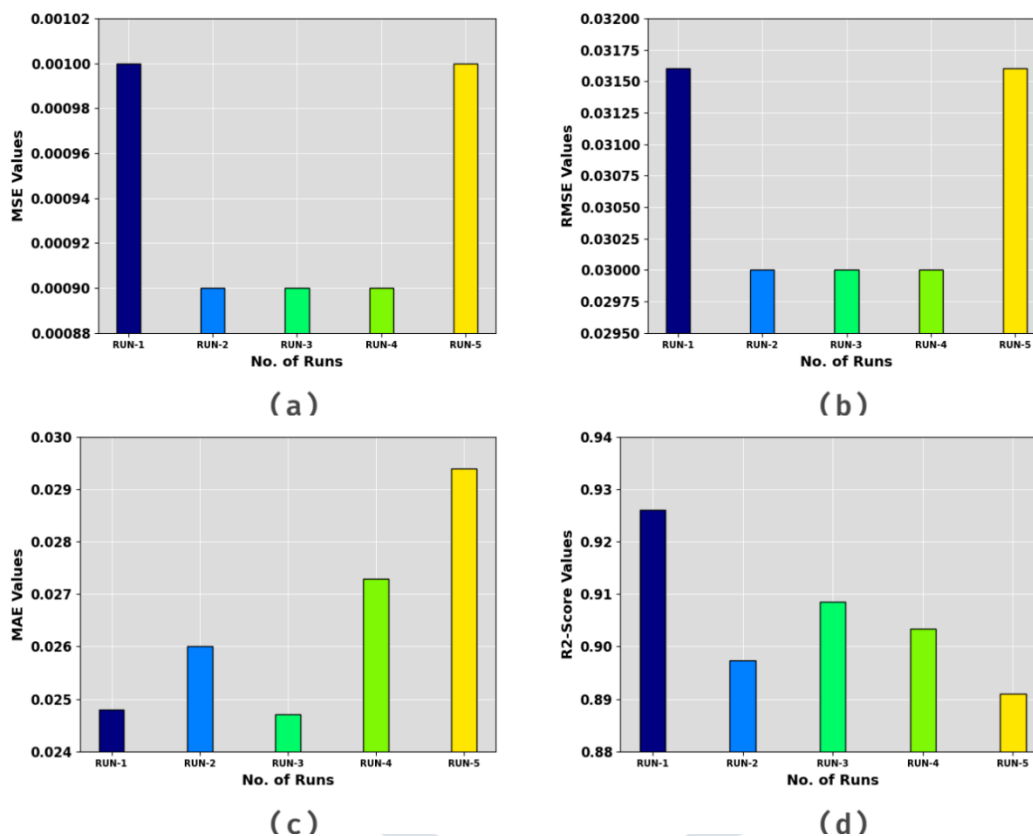
In which, TP and FP implies the true and false positive rates.

4. Performance Validation

This section inspects the predictive results of the MFOEML-RCYP methodology under several runs. In Table 1 and Fig. 2, the overall prediction outcome of the MFOEML-RCYP system is given. The outcomes appeared that the MFOEML-RCYP system obtains effective performance. Based on MSE, the MFOEML-RCYP technique offers reduced MSE of 0.0010, 0.0009, 0.0009, 0.0009, and 0.0010 under runs 1-5, correspondingly.

Table 1 Prediction outcome of MFOEML-RCYP technique with five runs

No. of Runs	MSE	RMSE	MAE	R2-Score
RUN1	0.0010	0.0316	0.0248	0.9261
RUN2	0.0009	0.0300	0.0260	0.8974
RUN3	0.0009	0.0300	0.0247	0.9085
RUN4	0.0009	0.0300	0.0273	0.9034
RUN5	0.0010	0.0316	0.0294	0.8910
AVERAGE	0.0009	0.0306	0.0264	0.9053

**Fig. 2.** Prediction outcome of a) MSE b) RMSE c) MAE d) R2-Score

Also, based on RMSE, the MFOEML-RCYP algorithm obtains lesser RMSE of 0.0316, 0.0300, 0.0300, 0.0300, and 0.0316 under runs 1-5, correspondingly. Besides, based on MAE, the MFOEML-RCYP methodology offers minimal MAE of 0.0248, 0.0260, 0.0247, 0.0273, and 0.0294 under runs 1-5, correspondingly. Moreover, based on R2-score, the MFOEML-RCYP approach attains lower R2-score of 0.9261, 0.8974, 0.9085, 0.9034, and 0.8910 under runs 1-5, correspondingly.

Fig. 3 represents the overall average analysis of MFOEML-RCYP technique under five runs with different measures. The experimental values implies that the MFOEML-RCYP algorithm has outperformed average MSE of 0.0009, RMSE of 0.0306, MAE of 0.0264, and R2-score of 0.9053, respectively.

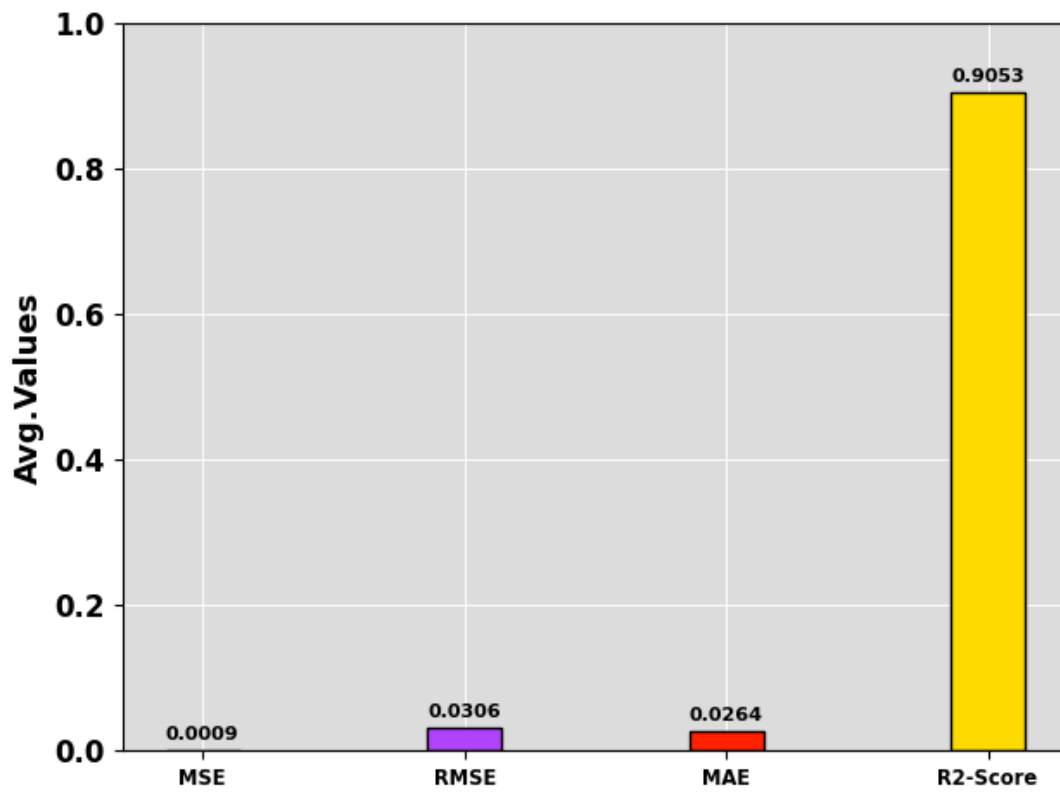


Fig. 3. Average of MFOEML-RCYP technique with different measures

The comprehensive MSE outcomes of the MFOEML-RCYP technique with comparison studies are reported in Table 2 and Fig. 4 [22]. The outcomes detected that the MFOEML-RCYP algorithm gains least MSE of 0.009. On the other hand, the TLBODL-RCYP, ANN-LM, ANN-BR, and ANN-SCG approaches attain increasing MSE of 0.0013, 0.0400, 0.0930, and 0.0840, correspondingly. Therefore, the MFOEML-RCYP methodology can be executed for accurate RCYP process.

Table 2 MSE outcome of MFOEML-RCYP technique with other approaches

Methods	MSE
MFOEML-RCYP	0.0009
TLBODL-RCYP	0.0013
ANN-LM	0.0400
ANN-BR	0.0930
ANN-SCG	0.0840

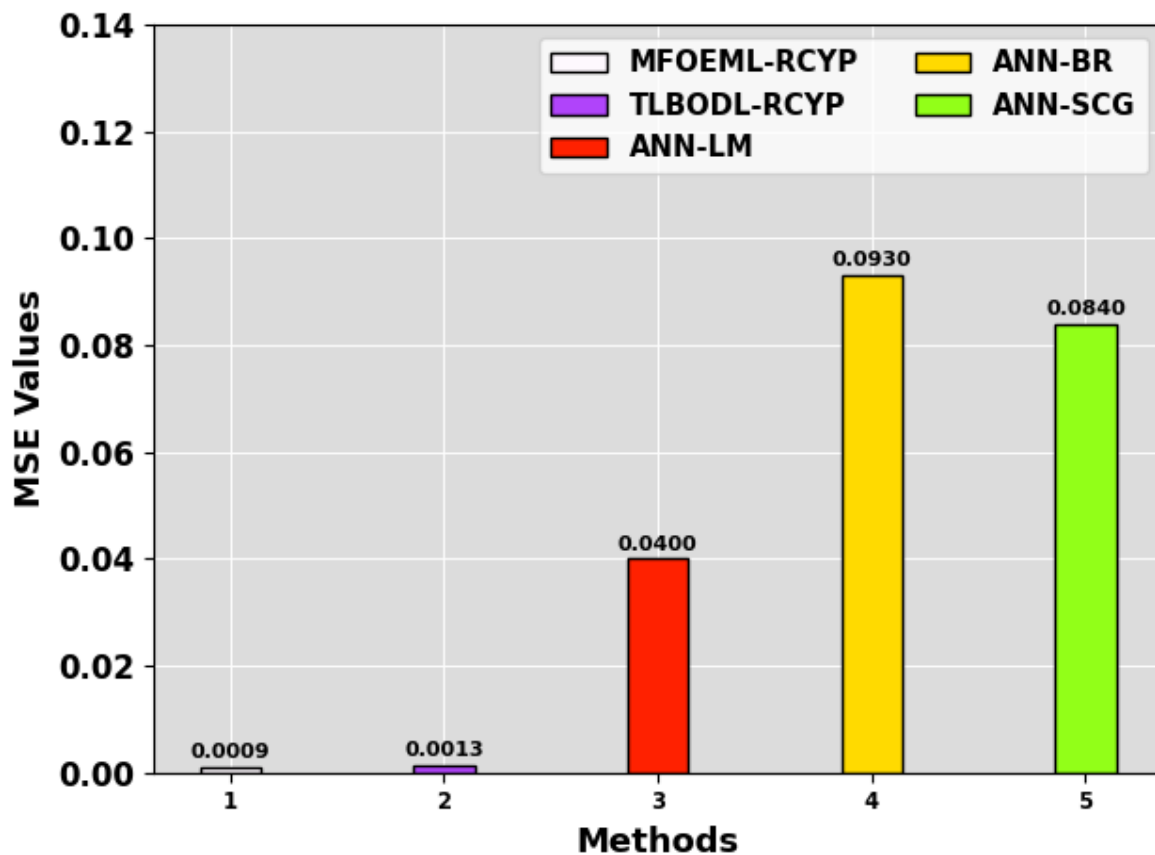


Fig. 4. MSE outcome of MFOEML-RCYP technique with other approaches

5. Conclusion

In this study, a novel MFOEML-RCYP approach is presented. The main drive of the MFOEML-RCYP model comprises three different kinds of procedures namely Z-score normalization, ensemble learning, and parameter optimizer. Initially, the MFOEML-RCYP technique exploits Z-score normalization to pre-process input characteristics, which ensures optimum data dissemination for succeeding analysis. To optimize predictive accuracy, an ensemble technique is introduced, integrating the strengths of various techniques such as BPNN, WNN, and Light GBM. Moreover, parameter optimization is implemented by the MFO for maximizing predictive outcomes and adjusting model parameters. Empirical outcomes illustrate the effectiveness of the presented MFOEML-RCYP technique in precisely predicting rice crop yields, emphasizing its prospective for enhancing crop management approaches and informing agricultural decision-making.

References

- [1] Kuradusenge, M., Hitimana, E., Hanyurwimfura, D., Rukundo, P., Mtonga, K., Mukasine, A., Uwitonze, C., Ngabonziza, J. and Uwamahoro, A., 2023. Crop Yield Prediction Using Machine Learning Models: Case of Irish Potato and Maize. *Agriculture*, 13(1), p.225.
- [2] Cao, J., Zhang, Z., Tao, F., Zhang, L., Luo, Y., Zhang, J., Han, J. and Xie, J., 2021. Integrating multi-source data for rice yield prediction across China using machine learning and deep learning approaches. *Agricultural and Forest Meteorology*, 297, p.108275.

- [3] Abbas, F., Afzaal, H., Farooque, A.A. and Tang, S., 2020. Crop yield prediction through proximal sensing and machine learning algorithms. *Agronomy*, 10(7), p.1046.
- [4] Zhou, X., Kono, Y., Win, A., Matsui, T. and Tanaka, T.S., 2021. Predicting within-field variability in grain yield and protein content of winter wheat using UAV-based multispectral imagery and machine learning approaches. *Plant Production Science*, 24(2), pp.137-151.
- [5] Chandra, A., Mitra, P., Dubey, S.K. and Ray, S.S., 2019. Machine learning approach for kharif rice yield prediction integrating multi-temporal vegetation indices and weather and non-weather variables. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, pp.187-194.
- [6] Medar, R.A., Rajpurohit, V.S. and Ambekar, A.M., 2019. Sugarcane crop yield forecasting model using supervised machine learning. *International Journal of Intelligent Systems and Applications*, 11(8), p.11.
- [7] Bondre, D.A. and Mahagaonkar, S., 2019. Prediction of crop yield and fertilizer recommendation using machine learning algorithms. *International Journal of Engineering Applied Sciences and Technology*, 4(5), pp.371-376.
- [8] Nosratabadi, S., Ardabili, S., Lakner, Z., Mako, C. and Mosavi, A., 2021. Prediction of food production using machine learning algorithms of multilayer perceptron and ANFIS. *Agriculture*, 11(5), p.408
- [9] Son, N.T., Chen, C.F., Cheng, Y.S., Toscano, P., Chen, C.R., Chen, S.L., Tseng, K.H., Syu, C.H., Guo, H.Y. and Zhang, Y.T., 2022. Field-scale rice yield prediction from Sentinel-2 monthly image composites using machine learning algorithms. *Ecological informatics*, 69, p.101618.
- [10] Shobha, M.S., Hegde, S.V., Yashwanth, C.V. and Kiran, S.C., 2020. Crop Yield Production using Machine Learning Algorithm. *International Research Journal of Engineering and Technology (IRJET)*, 4(7), pp.3328-3330.
- [11] Hoque, M.J., Islam, M.S., Uddin, J., Samad, M.A., De Abajo, B.S., Vargas, D.L.R. and Ashraf, I., 2024. Incorporating Meteorological Data and Pesticide Information to Forecast Crop Yields Using Machine Learning. *IEEE Access*.
- [12] Kilaru, R. and Raju, K.M., 2022, February. Prediction of maize leaf disease detection to improve crop yield using machine learning based models. In *2021 4th International Conference on Recent Trends in Computer Science and Technology (ICRTCST)* (pp. 212-217). IEEE.
- [13] Muthukumaran, S., Geetha, P. and Ramaraj, E., 2023. Multi-Objective Optimization with Artificial Neural Network Based Robust Paddy Yield Prediction Model. *Intelligent Automation & Soft Computing*, 35(1).
- [14] Ramisetty, S., Bansode, D., Atmakur, V.K., Ghantasala, G.P., Ushasree, D. and Kumar, M., 2024. Zone-II & III: Machine Learning based Rice Yield Prediction in Andhra Pradesh. In *MATEC Web of Conferences* (Vol. 392, p. 01086). EDP Sciences.
- [15] Reyana, A., Kautish, S., Karthik, P.S., Al-Baltah, I.A., Jasser, M.B. and Mohamed, A.W., 2023. Accelerating Crop Yield: Multisensor Data Fusion and Machine Learning for Agriculture Text Classification. *IEEE Access*, 11, pp.20795-20805.
- [16] Devi, B.S., Sandhya, N. and Chatrapati, K.S., 2024. Hybrid deep WaveNet-LSTM architecture for crop yield prediction. *Multimedia Tools and Applications*, 83(7), pp.19161-19179.

- [17] Imron, M.A. and Prasetyo, B., 2020. Improving algorithm accuracy k-nearest neighbor using z-score normalization and particle swarm optimization to predict customer churn. *Journal of Soft Computing Exploration*, 1(1), pp.56-62.
- [18] Bao, T., Chen, K., Zhang, H., Zhang, Z., Ai, Q. and Yan, J., 2023. Comprehensive Calculation Method of Semantic Similarity of Transport Infrastructure Ontology Concept Based on SHO-BP Algorithm. *Applied Sciences*, 13(19), p.10587.
- [19] Yao, Q., Song, X. and Xie, W., 2024. State of health estimation of lithium-ion battery based on CNN–WNN–WLSTM. *Complex & Intelligent Systems*, pp.1-18.
- [20] Dasari, S. and Kaluri, R., 2024. An Effective Classification of DDoS Attacks in a Distributed Network by Adopting Hierarchical Machine Learning and Hyperparameters Optimization Techniques. *IEEE Access*.
- [21] Nandakumar, N. and Symon, V.T.P., 2024. An implementation of inertia control strategy for grid-connected solar system using moth-flame optimization algorithm. *Automatika*, 65(1), pp.206-216.
- [22] Amaratunga, V., Wickramasinghe, L., Perera, A., Jayasinghe, J. and Rathnayake, U., 2020. Artificial neural network to estimate the paddy yield prediction using climatic data. *Mathematical Problems in Engineering*, 2020.

