Agriculture crop yield prediction using inertia based cat swarm optimization

Dwaram Jayanarayana Reddy[1], Rudra Kumar Madapuri[2]
[1]Department of Computer Science and Engineering, Jawaharlal Nehru Technological University Anantapur, Ananthapuramu, India
[2]Department of Computer Science and Engineering, G Pullaiah College of Engineering and Technology (Autonomous), Kurnool, India

Article Info

Article history:

Received Jul 25, 2023
Revised Oct 11, 2023
Accepted Oct 20, 2023

Keywords:

Cat swarm optimization
Crop yield prediction
Cultivation
Recurrent neural network
Time series data

ABSTRACT

Crop yield prediction is among the most important and main sources of income in the Indian economy. In this paper, the improved cat swarm optimization (ICSO) based recurrent neural network (RNN) model is proposed for crop yield prediction using time series data. The inertia weight parameter is added to position equation that is selected randomly, and a new velocity equation is produced which enhances the searching ability in the best cat area. By using inertia weight, the ICSO enhances performance of feature selection and obtains better convergence in minimum iteration. The RNN is applied to produce direct graph using sequence of data and decides current layer output by involving all other existing calculations. The performance of the model is estimated using coefficient of determination (R2), root mean square error (RMSE), mean squared error (MSE), and mean absolute error (MAE) on the yield from the years 2011 to 2021 with an annual prediction for 120 records of approximately 8 million nuts. The evaluated result shows that the proposed ICSO-RNN model delivers metrics such as R2, MAE, MSE, and RMSE values of 0.99, 0.77, 0.68, and 0.82 correspondingly, which ensures accurate yield prediction when compared with the existing methods which are hybrid reinforcement learning-random forest (RL-RF) and machine learning (ML) methods.

Corresponding Author:

Dwaram Jayanarayana Reddy
Department of Computer Science and Engineering, Jawaharlal Nehru Technological University Anantapur
Ananthapuramu, India
Email: djnreddy@gmail.com

1. INTRODUCTION

Agriculture is one of the most important occupations in the development of the Indian economy. Every farmer is interested in knowing the crop yield that is expected during the harvest period [1]. The production of agriculture is needed to be improved to tackle the problem of food scarcity [2]. There are several reasons for crop loss; the most important factor is that the crop productivity and fertilizers are not accurately understood by the farmers [3]. The crop yield prediction along with the prediction of climate change will be useful for the farmers in cultivation. The major challenge of crop prediction is climate change as the weather decides the crops yield [4]. Over the last few years, various researchers have concentrated on the improvement of crop yield prediction through a variety of methods such as crop growth process-oriented methods and statistical empirical models [5]. The crop yield prediction is based on numerous environmental components like soil, temperature, humidity, fertilizers, rainfall, pH value and other parameters [6]. Among the various factors, soil plays an important role in agriculture to obtain good crop yields. There are several techniques followed by farmers that include composite, organic, and physical procedures for crop cultivation which help to manage the soil quality [7]. Crop yield prediction is performed using various techniques, but the most common approach is remote sensing. It produces reasonable temporal resolution and minimum-cost non-destructive information on crops [8]. Additionally, machine learning techniques based on remote sensing data are established for agriculture models [9]. Crop growth model, radiation use efficiency (RUE), and empirical methods are the three crop yield prediction techniques used in remote sensing. Crop growth model imitates the crop yield depending on remote sensors, while RUE evaluates the crop yield [10]. The empirical method is related to the yield which is directly removed by the remote sensors [11]. Machine learning techniques such as artificial neural network (ANN), support vector machine (SVM), decision tree (DT), random forest (RF), principle component analysis (PCA), clustering, and reinforcement learning (RL) are employed for crop yield prediction [12]. Among the several machine learning approaches, the RF and SVM obtain better results in predicting the crop yield. The advantage of utilizing machine learning (ML) techniques is that they decrease crop loss and help in achieving a high price [13]. Due to the data complexity, crop yield prediction is a difficult task for farmers to carry out effectively. There are issues in creating an empirical expert system through knowledge as it cannot be categorized as time-series data to predict crop yield [14]. The mathematical studies and historical data from the previous years are helpful in certain regions, but cannot be employed frequently [15]. Due to the limited resources, crop yield is utilized in remote areas where the meteorological data is unavailable, but it struggles to handle the temporal dependencies and patterns in time-series data [16].

Paudel *et al.* [17] implemented a machine learning technique for crop yield production using the data from European commission's MARS crop yield forecasting system (MCYFS). ML was a technique where huge numbers of data were gathered and joined with the agronomic principles of crop models to create a guideline for huge-scale crop yield forecasting. The guideline had ample room for development in both the fit-for-purpose optimizations and general design principles. Under this guideline, the data did not have a generic method for preprocessing, and big data analysis was not implemented. Elavarasan *et al.* [18] introduced a combination of correlation-based filter (CFS) and random forest-recursive feature elimination (RF-RFE) wrapper that depended on extracting features by predicting crop yield using a ML model. This model was utilized to identify the features of groundwater, climate and soil properties for building crop yield prediction. The benefit of using this hybrid filter model among the other existing filter models was that the computation time was shorter, and the result of this method gave better accuracy for predicting the crop yield. But, in this paper, eternal variables related to weed infestations and pesticides were not considered for crop yield prediction. Sun *et al.* [19] implemented a winter wheat yield prediction method depending on ML and multi-source data in China from 2014–2018. In this paper, the RF and multiple linear regression (MLR) models were employed to predict the crop yield by using climate, geographic and satellite data. The advantage of using this model was better accuracy, and a simple, inexpensive and scalable framework. Additionally, this model avoided the randomness of a single model. In this paper, the spatial distribution interannual information was not considered from satellite data to minimize the errors in crop yield prediction. Elavarasan and Vincent [20] introduced a deep reinforcement learning model for predicting crop yields using sustainable Agrarian applications. In this proposed deep reinforcement learning (DRL)-based method, the reinforcement agent was unable to appreciate the input data, and during the learning process, the samples were not effectively learned. In this model, the agent performed some actions and received a score by maximizing the prediction accuracy while minimizing the error value. In the DRL model, if the time series was much longer, then it caused the gradients to disappear or explode for predicting the crop yield. Rajagopal *et al.* [21] developed a technique for predicting crop yield through recurrent cuckoo search optimization (RCSO) neural networks. The developed artificial intelligent (AI) system combined the several factors like rainfall, soil and crop yield production to forecast market price. Feature selection was done by following the feature extraction approach. The better features were optimized using the developed algorithm, then the output was given as input for classification process which was done by using a discrete deep belief network (DBN)-very deep convolutional networks (VGGNet) classifier. The advantage of using this model was yielding the right crop at a given time, economic growth, balancing of crop production and minimizing scarcity. However, the proposed model only performed well in theoretical aspects and not in practical aspects. Gupta *et al.* [22] introduced a ML and feature selection for early and precise crop yield prediction. The dataset included all information related to crops which were collected, and then the relief algorithm was used for selecting the features. This selection was useful for achieving exact result by classifying significant characteristics related to a specific real-world problem. The linear discriminant analysis (LDA) algorithm was used for extracting the features. The ML algorithms like particle swarm optimization (PSO)-support vector machine (SVM), random forest (RF) and K-nearest neighbor (KNN) were used for classifying the crop yield. This method was most beneficial for farmers and policymakers to be able to accurately predict crop yields through the successive seasons.

Ali *et al.* [23] implemented an online sequential extreme learning machine model (OSELM) with ant colony optimization (ACO) for wheat yield prediction. The developed dual-phase hybrid ACO-OSELM was estimated through ACO-RF and ELM methods using metrics. A parameter called pheromone in ACO was allocated for predictor stations that classified the predictors along with test or target station at the stating stage. The developed dual-phase model provided better accuracy with minimum error statistics. Due to the limited resources, the developed methodology was utilized in river areas where meteorological data was unavailable. Elavarasan and Vincent [24] presented a reinforced random forest model for improving crop yield prediction by incorporating agrarian parameters. A reinforcement learning (RL) process enclosed in RF was equipped at every node to decide the measures. This model had a good learning speed with minimum configurations and offered a better optimization framework. The model's behavior was sensitive to modifications in reward functions. The drawback of this method was the excessive baseline function which was used to achieve tight error bound. Batool *et al.* [25] implemented a hybrid technique for tea crop yield prediction by employing the machine learning model and simulation model using the food and agriculture organization (FAO) of the United Nations' AquaCrop crop growth model. This hybrid proposed method used data on soil, weather and agriculture management from the national tea and high-value crop research institute (NTHRI). By comparing both the model results, it is seen that the ML model performed better than the simulation models. But this paper did not consider the superlative hybrid among the four machine learning methods, RF, RL, SVM and DT that were carried out in the performance evaluation metrics. From the overall analysis, the existing methods are seen to have the limitations such as struggles to handle temporal dependencies and patterns in time-series data, data overfitting issues, good outcomes only in theoretical aspects and not in practical aspects, and being utilized in river areas due to limited resources where the meteorological data was unavailable which is crucial in crop yield prediction. The proposed improved cat swarm optimization (ICSO) based recurrent neural network (RNN) model overcomes these limitations. It enhances the outcome and obtains better convergence in minimum iteration. The inertia weight parameter is added to the position equation that will be selected randomly, and a new velocity equation is produced which enhances the searching ability in the best cat area. By using inertia weight, the ICSO is deployed to enhance the performance in feature selection. The RNN is applied to produce a direct graph by using sequence of data, and it decides the present layer output by involving all other existing calculations.

The major contribution of this research is mentioned as: i) The dataset is collected from Andhra Pradesh and it consists of the monthly yield of crops from 10 districts from the period of 2011-2021, with an annual production for 120 records of 8 million nuts approximately. ii) By using inertia weight, the ICSO is utilized to enhance the performance of feature selection and obtains better convergence in minimum iteration. And iii) The RNN is utilized for classification that produces direct graph using sequence of data and decides present layer output by involving all other existing calculations.

The rest of the manuscript is organized as follows: section 2 illustrates the proposed method. The improved cat swarm optimization-based feature selection is presented in section 3. The results and discussion are illustrated in section 4. Section 5 describes the conclusion of this paper.

2. PROPOSED METHOD

This proposed methodology is developed for predicting the agriculture crop yield. The inertia weight parameter is added to the position equation that is selected randomly, while a new velocity equation is produced which enhances the searching ability in the best cat area. By using inertia weight, the ICSO is utilized to enhance the performance of feature selection and obtains better convergence in minimum iteration. In this paper, Figure 1 illustrates four different steps that include data collection, preprocessing, feature selection, classification.

2.1. Data collection

The dataset is collected from Andhra Pradesh and it consists of the monthly yield of crops from 10 districts namely, Kakinada, Anakapalli, East Godavari, West Godavari, Guntur, Eluru, Prakasam, Tirupati, Chittoor and Visakhapatnam from the period of 2011-2021 with annual production for 120 records of approximately 8 million nuts. The features such as region, soil, season, year, tones, crop, acres, water, area, temperature, production, solar radiation and pH value are taken into consideration. This dataset can be further split into two parts: 80% for training set and 20% for testing set. After collecting the dataset, data is preprocessed by using the min-max normalization method.

2.2. Preprocessing

Data preprocessing is a process of converting raw data into a desired format as the dataset from several resources may consist of incomplete data. So, for further analysis, this data requires to be filtered and normalized [26]. Data normalization is the process of preprocessing the input data. Andhra Pradesh District dataset is standardized using min-max normalization method. The highest score of the feature is transformed into 1, the smallest score of the feature is transformed into 0, and other values of the feature are transformed into an integer between 0 and 1. The mathematical representation of min-max normalization [27] is shown in (1).

$$v' = \frac{v - min_A}{max_A - min_A}\left(new_{max_A} - new_{min_A}\right) + new_{min_A} \quad (1)$$

where, v' is the respective min-max normalization value of the attribute, v is the primary value of the attribute, new_{max_A} is the new maximum score of the feature, new_{min_A} is the new smallest score of the feature. After the preprocessing step, the features are selected by using optimization algorithm.

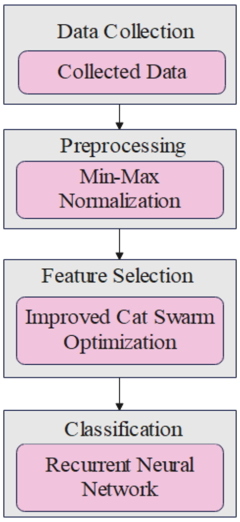

Figure 1. Proposed method for crop yield prediction

3. IMPROVED CAT SWARM OPTIMIZATION BASED FEATURE SELECTION

In this proposed method, an improved cat swarm optimization algorithm [28] is used to select a count of six features which are, soil, temperature, water, production, solar radiation and area. In this ICSO, first how many cats to be used is chosen, and then the cats within the ICSO model are implemented to resolve the optimization issues. The behaviors of the cat are classified into two modes as, seeking mode and tracing mode. The ICSO obtains better performance at the final output, until it reaches the iterations at the end. A deep learning-based methodology is proposed for predicting agriculture crop yield. The inertia weight parameter is added to the position equation that is selected randomly and a new velocity equation is produced to enhance the searching ability in the best cat area. By using inertia weight, the ICSO is utilized to enhance the outcome of feature selection and classification, and hence obtain better convergence in minimum iteration.

3.1. Seeking mode

The seeking mode is employed for processing the cat which moves carefully and slowly. In the seeking mode, there are four necessary elements which are, seeking memory pool (SMP), seeking range of selected dimension (SRD), counts dimensions to change (CDC) and self-position consideration (SPC). The SMP is employed to decide the measurements of an individual cat and it represents the desired position of the cat. From the memory pool, the cat selects a position by using the corresponding rules. The SRD is specified as the selected dimension of the mutative ratio, and it has a condition for transforming the size measurement. In the seeking mode, the measurement is chosen for mutating where the comparison between the existing score and the new score is within a specified limit. The CDC exhibits the value of how many size measurements differ. In the seeking mode, these factors play a significant role. The SPC determines the cat’s standing areas which are the moving positions of the candidate. The SPC score is considered as a Boolean value, either true or false. The procedures of the seeking mode are represented:

a. Consider the current position in j number of copies for cat_k, here $j = SMP$ if the SPC score is considered as true, and $j = SMP - 1$ if SPC score is considered as false. Then, the present position is considered as an individual candidate solution.

b. For all the copies that depend on CDC, SRD process of the present value is minimized or maximized, and then the existing value is replaced as presented in (2),

$$Xjd_{new} = (1 + rand \times SRD) \times Xjd_{old} \tag{2}$$

where, Xjd_{new} and Xjd_{old} are the updated position and the existing position of cat_k, respectively. d represents the dimensions, j illustrates the total amount of cats and $rand$ is the random number between the range of [0,1].
c. The fitness value of the overall candidates is estimated.
d. If the entire fitness value is not accurately the same, the selecting possibility of all the candidates is evaluated using the below formula, or else the selecting possibility of all the candidates is to be fixed to 1.
e. Selecting the random position moving from candidate point and replacing the position of cat_k is formulated as shown in (3),

$$P_i = \frac{|FS_i - FS_b|}{FS_{max} - FS_{min}} \tag{3}$$

If the goal of the fitness score is to reduce, then $FS_b = FS_{max}$, or else $FS_b = FS_{min}$, where the FS_{min} and FS_{max} represents the smallest and highest candidate fitness scores, correspondingly.
f. The candidates are classified and selected by P_i and the position of cat_k is replaced.

3.2. Tracing mode

The tracing mode is a process of tracing some targets. Once the cat is within the trace mode, it shifts every dimension to its corresponding speed. The procedure of tracing mode is represented in the following steps:

Step 1: The velocity of all dimensions $(v_{k,d})$ are updated corresponding to (4),

$$V_{k,d} = V_{k,d} + r_1 c_1 (X_{best,d} - X_{k,d}) \tag{4}$$

where, $V_{k,d}$ is the velocity of the cat_k in d dimension, r_1 and c_1 illustrate the random score in the range of [0,1] and the fixed value, respectively. $X_{k,d}$ is a coordination of cat_k in d dimension, $X_{best,d}$ is the coordination with the better result in the dimension d.

Step 2: It is verified that the velocity of each cat is adjacent to the highest speed. In that instance, the updated speed score is in the high range, and the updated score is set equivalent to the restricted limit.

Step 3: The cat's position is updated by inserting the updated speed and then the entire cat's position is gained as formulated in (5),

$$X_{k,d} = X_{k,d} + V_{k,d} \tag{5}$$

where, $V_{k,d}$ illustrates velocity of the cat_k in the d^{th} dimension, $X_{k,d}$ illustrates the coordination of the cat_k in the d^{th} dimension. Then the fitness score of updated cat position of entire cat is estimated. At that time, the tracing mode and searching mode are merged by employing a mixture ratio. Most of the time, the cats rest, and the remaining time, they move slowly and carefully. As the cat spends a huge time in the searching mode, it is concluded that the mixture ratio is a minimum value.

In this paper, the ICSO is proposed to enhance the performance of feature selection and to obtain better convergence in minimum iteration. The inertia weight parameter is added to the position equation that is selected randomly, and a new velocity equation is produced to enhance the searching ability in the best cat area. By using this parameter, the global and local search ability are balanced. A maximum inertia weight simplifies the global search, while a minimum inertia weight simplifies the local search. Initially, the high value is utilized and then is gradually lowered to the minimum value. For every iteration, maximum inertia weight occurs at the first dimension which is then decreasingly updated in each dimension, and the velocity is updated to a new form that is modifiable in every cat. The computational process of ICSO is represented below:

Step 1: Every parameter like velocity, cat position, population, SRD, SMP, structure of neighborhood, α, β and c is initialized.

Step 2: N number of cats are located in random search space.

Step 3: Cat fitness function (FF) is calculated and the cat's best position is stored in memory.

Step 4: Based on MR, the flag is allocated and corresponding to flag value, and then the cats are randomly distributed into the tracing mode (TM) and seeking mode (SM).
Step 5: $If\ (Flag == 1)$, then the cat is in SM.
Step 6: k copies of every cat are made.
Step 7: The bit score for all cats are evaluated through SRD and every cat is added or removed to/from the shifting scores.
Step 8: The FF of all new cat positions is evaluated and the cat's best position is stored in memory.
Step 9: Else, cat is in TM.
Step 10: The velocity and position of all cats are updated by (4) and (5).
Step 11: The FF of newly produced cat's position is evaluated and the cat's best position is stored in memory.
Step 12: The cat's position is updated and also the best position of cat is decided.
Step 13: The final solutions are achieved.
The ICSO provides key points according to the CSO i.e., the cat position and best key points are selected as the best cat position.

3.3. Fitness function

In the proposed ICSO algorithm, fitness function is employed for estimating the fitness of cat. To solve the imbalance issues in the optimization algorithm, the frequently used fitness function is taken as the weighted average classification accuracy (avg) for classifying the imbalanced data. The fitness function is mandatory to give the weight, but also the assurance of weight is specific. To commonly treat the majority and minority class, the weight w= 0.5 is used. In imbalanced data classification, the area under curve (AUC) and Pearson correlation are significant measures and are also used in fitness functions. In common, glenohumeral joint osteoarthritis (GJO) with AUC as a fitness function consumes large training time and gives better outcomes in evaluation. AUC is represented in (6), (7) and (8),

$$AUC_F = \sum_{I=1}^{N-1} \frac{1}{2} \times (FPR_{i+1} - FPR_i)(TPR_{i+1} - TPR_i) \quad (6)$$

where, N is the threshold number, FPR_i and TPR_i are the i^{th} threshold of false positive rate and true positive rate respectively.

$$AUC_w = \frac{\sum_{i \in Min} \sum_{j \in Maj} I_{wmw}\,(P_i, P_j)}{|Min| * |Maj|} \quad (7)$$

where,

$$I_{wmw}(P_i, P_j) = \begin{cases} 1, & P_i > P_j \ and\ P_i \geq 0 \\ 0, & otherwise \end{cases} \quad (8)$$

where, P_i and P_j represent the output values taking an instance of i and j from the minority class and majority class as input respectively. $|Min|$ and $|Maj|$ illustrate the number of minority class and majority class instances, respectively. Pearson correlation is formulated in (9),

$$r = \frac{N\sum xy - (\sum x)(\sum y)}{\sqrt{N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}} \quad (9)$$

where r is the Pearson correlation coefficient, N is the pair scores count, $\sum xy$ is the sum and product of scores, $\sum x$ and $\sum y$ is the sum of x and y scores, $\sum x^2$ and $\sum y^2$ is the sum of squares of *x* and *y* scores.

3.4. Classification

After selecting the features from the crops, the recurrent neural network (RNN) is utilized for classification. The RNN is one category of neural network that produces a direct graph by using a sequence of data [29]. RNN decides the current layer output by involving all other existing calculations. It differs from a classic neural network, in which all the input values are autonomous to each other. The following steps constitute the training flow of RNN in classification.
Step 1: The network is supplied with an individual time of input.
Step 2: The present input and earlier state are used for evaluating the present state.
Step 3: For the upcoming time steps, the present state ht is transferred to the $ht - 1$.
Step 4: All times steps are built based on the problems. The instruction is merged from the earlier state.
Step 5: After completing all the time steps, the last present state is employed by the outputs.

Step 6: For updating the weights, the fault is back-propagated to the network. Therefore, RNN is trained.

In RNN, every input is transferred to the next layer while the output of every layer fully depends on the previous layer. Hence it is called a recurrent neural network, and it accomplishes the same function for every component. The function of RNN with time-series data [30] and the output of RNN obtain the best result which is utilized to determine the earlier information. The RNN combines the output and forgets the gate into a specific update gate up_d, in which the linear interpolation method is used to obtain a better result. Assume that the $e_d \leftarrow Fr_{k^*}$ is the input feature of dimension d, then H_{d-1} represents the hidden previous state. Equations (10) and (11) represent the output of the update gate up_d and reset gate rs_d.

$$up_d = acv(WF^{eup}e_d + WF^{Hup}H_{d-1}) \tag{10}$$

$$rs_d = acv(WF^{ers}e_d + WF^{Hrs}H_{d-1}) \tag{11}$$

where, acv represents the activation function, in general, it is a logistic sigmoid function. $WF^d = WF^{eup}, WF^{ers}, WF^{Hup}, WF^{Hrs}$ is denoted by the weight function which is tuned through training algorithm to decrease error variation among actual and forecasted outputs. The hidden unit candidate state represented in (12).

$$\overline{H_d} = tan\left(WF^{eH}e_d + WF^{HH}(H_{d-1} \otimes rs_d)\right) \tag{12}$$

where, $\otimes$ is the element-wise multiplication. The linear interpolation between candidate state $\overline{H_d}$ and H_{d-1} Additionally, the hidden activation function in the term d is represented as H_d of RNN as illustrated in (13).

$$H_d = (1 - up_d) \otimes \overline{H_d} + up_d \otimes H_{d-1}H_d = (1 - up_d) \otimes \overline{H_d} + up_d \otimes H_{d-1} \tag{13}$$

where, H_{d-1} denotes the hidden previous state and $\otimes$ is the element-wise multiplication. In neural networks, all output and input values are independent, whereas in the RNN, the data in the hidden layer is remembered and it is completed in a previous hidden layer. The remembered data is permitted as input to the next layer so that it can perform operations on the data. This procedure is continuously repeated until it obtains the preferred output which is formulated in (14), (15), and (16),

$$h_t = f(h_{t-1}, x_t) \tag{14}$$

$$h_1 = tan\,h\ (W_{hh}h_{t-1} + W_{xh}X_t) \tag{15}$$

$$y_t = W_{hv}h_t \tag{16}$$

This RNN takes the sequence of inputs from x_0 and it outputs the sequence as h_0 which combines with the $x_1$which is the input for the next process. Then, the x_1 and h_0 are inputs for next process and similarly, h_1 and x_2 are inputs for the next step. By selecting six different features of soil, temperature, water, production, solar radiation and area, the RNN produces a direct graph using the sequence of data and decides present layer output by involving all other existing calculations.

The proposed RNN model contains advantages like enhancing crop yields, real-time crop analysis, choosing effective parameters, and making smart decisions. Additionally, it helps to reduce the loss to the farmers when unfavorable conditions occur. The RNN performs well in time series forecasting issues like crop yield prediction. It can capture long-term dependencies in the data which enables it to learn complex relationships and patterns. It also learns relevant features automatically from the input data. Due to the ability to capture complex relationships and patterns, RNN can yield higher prediction accuracy compared to traditional statistical models.

4. RESULTS AND DISCUSSION

In this paper, the proposed model is simulated through python environment with the system requirements of RAM 16 GB, Intel core i7 processor, and a Windows 10 (64 bit) operating system. Statistical parameters like the coefficient of determination (R2), root mean square error (RMSE), mean squared error (MSE), and mean absolute error (MAE) are used for estimating the model. These parameters are calculated from the predicted yields of all the years for each district. The mathematical representation of these parameters is shown in (17) to (20),

$$R^2 = 1 - \sum_{i=1}^{n} \left(\frac{(y_i - y_{di})^2}{(y_{di} - y_m)^2} \right) \tag{17}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - y_{di}| \tag{18}$$

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - y_{di})^2 \tag{19}$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - y_{di})^2} \tag{20}$$

where, n is number of points, y_i is predicted value obtained from neural network, y_{di} is real value, and y_m is mean of the real value.

4.1. Quantitative analysis

This section shows the quantitative analysis of the ICSO-RNN model without augmentation in R2, MSE, MAE, and RMSE with actual and optimized features. Tables 1 and 2 illustrate the quantitative analysis of various classifiers with actual and optimized features. Table 3 illustrates the quantitative analysis of various optimization algorithms. The features such as region, soil, season, year, tones, crop, acres, water, area, temperature, production, solar radiation and pH value are considered. Out of these 13 features, only 6 features are extracted, the remaining 7 features contain contributions which enhance the model's performance. The model's performance decreases because it does not extract seven features.

The performance of classifiers with actual features is shown in Table 1 and Figure 2. The R2, MSE, MAE, RMSE of ANN, convolutional neural network (CNN), long short-term memory (LSTM) and RNN are measured and compared with the proposed ICSO-RNN. The obtained result displays that the proposed ICSO-RNN model achieves a better result with the R2, MAE, MSE, and RMSE values of 0.95, 0.75, 0.76 and 0.87, respectively while considering feature selection compared to other classifiers.

The performance of classifiers with optimized features is compared to that with actual features is shown in Table 2 and Figure 3. The R2, MSE, MAE, RMSE of ANN, CNN, LSTM and RNN are measured and compared with the proposed ICSO-RNN. The obtained results show that the proposed ICSO-RNN model exhibits better performance with the R2, MAE, MSE and RMSE values of 0.99, 0.77, 0.68 and 0.82, respectively while considering feature selection compared to other classifiers.

Table 3 and Figure 4 describe the performance of RNN using various optimization techniques like PSO, genetic algorithm (GA), and ICSO. The obtained results display that the proposed ICSO-RNN achieves better results with the R2, MAE, MSE and RMSE values of about 0.99, 0.77, 0.68 and 0.82, respectively while evaluating feature selection compared to another optimization algorithms.

Table 1. Performance of the various classifier with actual features

Methods	R2	MAE	MSE	RMSE
ANN	0.70	0.89	0.88	0.93
CNN	0.85	0.85	0.85	0.92
LSTM	0.89	0.79	0.81	0.9
RNN	0.92	0.77	0.79	0.88
ICSO-RNN	0.95	0.75	0.76	0.87

Table 2. Performance of the various classifier with optimized features

Methods	R2	MAE	MSE	RMSE
ANN	0.75	0.87	0.85	0.92
CNN	0.89	0.84	0.81	0.9
LSTM	0.92	0.81	0.77	0.87
RNN	0.97	0.79	0.72	0.84
ICSO-RNN	0.99	0.77	0.68	0.82

Table 3. Performance of the proposed methodology

Methods	R2	MAE	MSE	RMSE
RNN	0.87	0.84	0.85	0.92
PSO-RNN	0.95	0.81	0.82	0.9
GA-RNN	0.98	0.79	0.79	0.88
ICSO-RNN	0.99	0.77	0.68	0.82

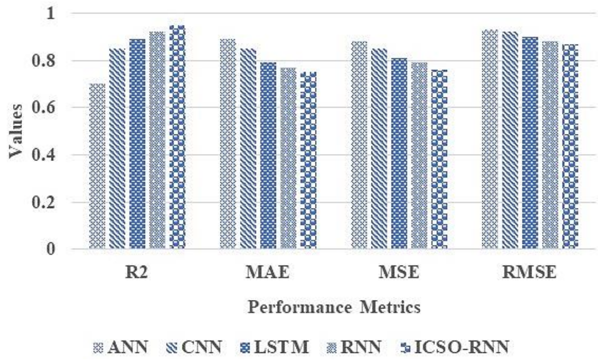

Figure 2. The performance of the classifier with actual features

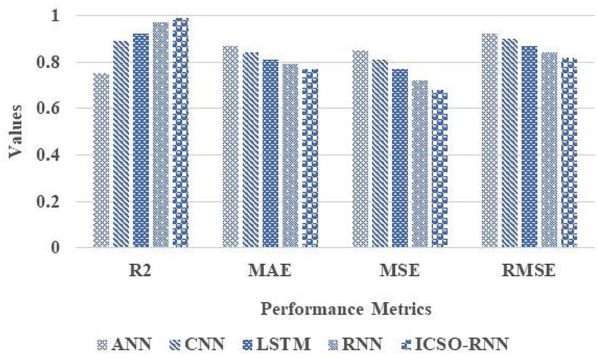

Figure 3. The performance of the classifier with optimized features

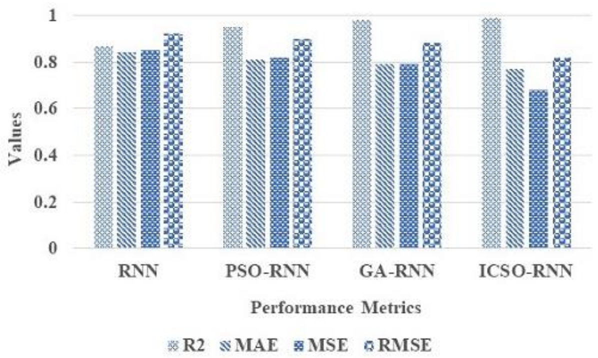

Figure 4. The performance of the proposed methodology

4.2. Comparative analysis

This section shows the comparison of RNN classifier in terms of R2, MSE, MAE, and RMSE as shown in Table 4. Existing researches such as ACO-OSELM [23], hybrid RF-RL [24] and ML [25] are used for evaluating the efficiency of the classifier. The proposed model is trained, tested and verified by applying a time series data. The proposed ICSO is evaluated for efficient classification by overcoming overfitting problems and classification by using the RNN technique.

4.2.1. Discussion

This section illustrates the proposed method's advantages and the limitations of the existing methods. The existing models such as the ACO-OSELM [23] model is utilized in remote areas where the meteorological data is unavailable because of limited resources. The hybrid RF-RL [24] model struggles to handle temporal dependencies and patterns in time-series data which is crucial for crop yield prediction. The inputs of the ML [25] model is extremely affected by sampling errors and limitations in data availability. For these reasons, the ICSO based RNN model is proposed to overcome these limitations. The introduced model enhances the outcomes and obtains better convergence in minimum iteration. The inertia weight parameter is added to the position equation that is selected randomly, and a new velocity equation is produced to enhance the searching ability in the best cat area. By using inertia weight, the ICSO is utilized to enhance the performance of feature selection, and hence obtains better convergence in minimum iteration. The RNN is applied to produce a direct graph by using sequence of data, additionally, it decides the present layer output by involving all other existing calculations.

Table 4. Comparative analysis of existing models with the proposed ICSO-RNN model

Author	Period	R2	MAE	MSE	RMSE
Ali *et al.* [23]	1981-2013	0.99	42.37	N/A	67.12
Elavarasan and Vincent [24]	2011-2016	0.87	0.17	0.06	0.23
Batool *et al.* [25]	2016-2019	N/A	0.12	0.02	0.15
Proposed ICSO-RNN	2011-2021	0.99	0.77	0.68	0.82

5. CONCLUSION

In this research, a new ICSO-RNN model is developed for effective agriculture crop yield prediction. This method contains multiple objective functions like min-max normalization and ICSO model for effective classification. The dataset is collected from Andhra Pradesh and it consists of the monthly yield of crops from 10 districts during the period of 2011-2021, with annual production for 120 records of approximately 8 million nuts. In min-max normalization, the highest score of the feature is transformed into 1, the smallest score of the feature is transformed into 0, and other values of the feature are transformed into an integer between 0 and 1. The inertia weight parameter is added to the position equation that is selected randomly, and a new velocity equation is produced to enhance the searching ability in the best cat area. By using inertia weight, the ICSO is utilized to enhance the performance of feature selection and to obtain better convergence in minimum iteration. RNN is utilized for producing a direct graph by using a sequence of data, and it decides the current layer output by involving all other existing calculations. The evaluated result shows that the proposed ICSO-RNN model delivers the evaluation metrics of R2, MAE, MSE, and RMSE with values as 0.99, 0.77, 0.68, and 0.82, respectively which ensures accurate yield prediction.

The proposed methodology has limitations such as its slow convergence in complex and high-dimensional problems, and generating complex and non-intuitive models, making it difficult to interpret. The data contains noise and uncertainties due to various factors. In time series data, it struggles to capture long-term dependencies. The future scope for the research is utilizing the hyperparameter tuning with deep learning algorithms like recurrent neural network for the better prediction of crop yield in time series data.

REFERENCES

[1] L. Li *et al.*, "Crop yield forecasting and associated optimum lead time analysis based on multi-source environmental data across China," *Agricultural and Forest Meteorology*, vol. 308–309, Oct. 2021, doi: 10.1016/j.agrformet.2021.108558.
[2] S. Shidnal, M. V. Latte, and A. Kapoor, "Crop yield prediction: two-tiered machine learning model approach," *International Journal of Information Technology*, vol. 13, no. 5, pp. 1983–1991, Oct. 2021, doi: 10.1007/s41870-019-00375-x.
[3] H. T. Pham, J. Awange, M. Kuhn, B. Van Nguyen, and L. K. Bui, "Enhancing crop yield prediction utilizing machine learning on satellite-based vegetation health indices," *Sensors*, vol. 22, no. 3, Jan. 2022, doi: 10.3390/s22030719.
[4] A. A. M. Ahmed, E. Sharma, S. J. J. Jui, R. C. Deo, T. Nguyen-Huy, and M. Ali, "Kernel ridge regression hybrid method for wheat yield prediction with satellite-derived predictors," *Remote Sensing*, vol. 14, no. 5, Feb. 2022, doi: 10.3390/rs14051136.
[5] Y. Guo *et al.*, "Integrated phenology and climate in rice yields prediction using machine learning methods," *Ecological Indicators*, vol. 120, Jan. 2021, doi: 10.1016/j.ecolind.2020.106935.
[6] M. G. Ziliani *et al.*, "Early season prediction of within-field crop yield variability by assimilating CubeSat data into a crop model," *Agricultural and Forest Meteorology*, vol. 313, Feb. 2022, doi: 10.1016/j.agrformet.2021.108736.
[7] J. Wang, H. Si, Z. Gao, and L. Shi, "Winter wheat yield prediction using an LSTM model from MODIS LAI products," *Agriculture*, vol. 12, no. 10, Oct. 2022, doi: 10.3390/agriculture12101707.
[8] M. F. de Oliveira, B. V. Ortiz, G. T. Morata, A.-F. Jiménez, G. de S. Rolim, and R. P. da Silva, "Training machine learning algorithms using remote sensing and topographic indices for corn yield prediction," *Remote Sensing*, vol. 14, no. 23, Dec. 2022, doi: 10.3390/rs14236171.
[9] V. Joshua, S. M. Priyadharson, and R. Kannadasan, "Exploration of machine learning approaches for paddy yield prediction in eastern part of Tamilnadu," *Agronomy*, vol. 11, no. 10, Oct. 2021, doi: 10.3390/agronomy11102068.
[10] Y. Zhao *et al.*, "The prediction of wheat yield in the North China plain by coupling crop model with machine learning algorithms," *Agriculture*, vol. 13, no. 1, Dec. 2022, doi: 10.3390/agriculture13010099.
[11] P. Dhandapani and A. Varadarajan, "Multi-channel convolutional neural network for prediction of leaf disease and soil properties," *International Journal of Intelligent Engineering and Systems*, vol. 15, no. 1, Feb. 2022, doi: 10.22266/ijies2022.0228.29.
[12] Y. Shen *et al.*, "Improving wheat yield prediction accuracy using LSTM-RF framework based on UAV thermal infrared and multispectral imagery," *Agriculture*, vol. 12, no. 6, p. 892, Jun. 2022, doi: 10.3390/agriculture12060892.
[13] S. Fei *et al.*, "UAV-based multi-sensor data fusion and machine learning algorithm for yield prediction in wheat," *Precision*

Agriculture, vol. 24, no. 1, pp. 187–212, Feb. 2023, doi: 10.1007/s11119-022-09938-8.
[14] I. Salehin *et al.*, "IFSG: Intelligence agriculture crop-pest detection system using IoT automation system," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 24, no. 2, pp. 1091–1099, Nov. 2021, doi: 10.11591/ijeecs.v24.i2.pp1091-1099.
[15] N. H. Huy Cuong *et al.*, "An approach based on deep learning that recommends fertilizers and pesticides for agriculture recommendation," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 5, Oct. 2022, doi: 10.11591/ijece.v12i5.pp5580-5588.
[16] M. Yessenova *et al.*, "The effectiveness of methods and algorithms for detecting and isolating factors that negatively affect the growth of crops," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 13, no. 2, pp. 1669–1679, Apr. 2023, doi: 10.11591/ijece.v13i2.pp1669-1679.
[17] D. Paudel *et al.*, "Machine learning for large-scale crop yield forecasting," *Agricultural Systems*, vol. 187, Feb. 2021, doi: 10.1016/j.agsy.2020.103016.
[18] D. Elavarasan, D. R. Vincent P M, K. Srinivasan, and C.-Y. Chang, "A hybrid CFS filter and RF-RFE wrapper-based feature extraction for enhanced agricultural crop yield prediction modeling," *Agriculture*, vol. 10, no. 9, Sep. 2020, doi: 10.3390/agriculture10090400.
[19] Y. Sun, S. Zhang, F. Tao, R. Aboelenein, and A. Amer, "Improving winter wheat yield forecasting based on multi-source data and machine learning," *Agriculture*, vol. 12, no. 5, Apr. 2022, doi: 10.3390/agriculture12050571.
[20] D. Elavarasan and P. M. D. Vincent, "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications," *IEEE Access*, vol. 8, pp. 86886–86901, 2020, doi: 10.1109/ACCESS.2020.2992480.
[21] A. Rajagopal, S. Jha, M. Khari, S. Ahmad, B. Alouffi, and A. Alharbi, "A novel approach in prediction of crop production using recurrent cuckoo search optimization neural networks," *Applied Sciences*, vol. 11, no. 21, Oct. 2021, doi: 10.3390/app11219816.
[22] S. Gupta *et al.*, "Machine learning- and feature selection-enabled framework for accurate crop yield prediction," *Journal of Food Quality*, vol. 2022, pp. 1–7, May 2022, doi: 10.1155/2022/6293985.
[23] M. Ali *et al.*, "Coupled online sequential extreme learning machine model with ant colony optimization algorithm for wheat yield prediction," *Scientific Reports*, vol. 12, no. 1, Mar. 2022, doi: 10.1038/s41598-022-09482-5.
[24] D. Elavarasan and P. M. D. R. Vincent, "A reinforced random forest model for enhanced crop yield prediction by integrating agrarian parameters," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 11, pp. 10009–10022, Nov. 2021, doi: 10.1007/s12652-020-02752-y.
[25] D. Batool *et al.*, "A hybrid approach to tea crop yield prediction using simulation models and machine learning," *Plants*, vol. 11, no. 15, Jul. 2022, doi: 10.3390/plants11151925.
[26] S. S. Mallela and S. K. Jonnalagadda, "Rainfall prediction based on spatial attention layer: a case study analysis," *International Journal of Intelligent Engineering and Systems*, vol. 15, no. 3, pp. 49–58, 2022, doi: 10.22266/ijies2022.0630.05.
[27] A. K. Ali and A. M. Abdullah, "Fake accounts detection on social media using stack ensemble system," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 3, pp. 3013–3022, Jun. 2022, doi: 10.11591/ijece.v12i3.pp3013-3022.
[28] H. Sikkandar and R. Thiyagarajan, "Deep learning based facial expression recognition using improved cat swarm optimization," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 2, pp. 3037–3053, Feb. 2021, doi: 10.1007/s12652-020-02463-4.
[29] V. Talasila, K. Madhubabu, M. Mahadasyam, N. Atchala, and L. Kande, "The prediction of diseases using rough set theory with recurrent neural network in big data analytics," *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 5, pp. 10–18, Oct. 2020, doi: 10.22266/ijies2020.1031.02.
[30] A. Muneer, R. F. Ali, A. Almaghthawi, S. M. Taib, A. Alghamdi, and E. A. A. Ghaleb, "Short term residential load forecasting using long short-term memory recurrent neural network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 5, pp. 5589–5599, Oct. 2022, doi: 10.11591/ijece.v12i5.pp5589-5599.

BIOGRAPHIES OF AUTHORS

Dwaram Jayanarayana Reddy is research scholar, Department of Computer Science and Engineering, JNTUA, Ananthapuramu-515002, Andhra Pradesh, India. Received M.Tech degree in computer networks and information security from JNTU Hyderabad, Andhra Pradesh, India. He is a life member of CSI, and he has published many research papers in national and international journals, and conferences. His research interests include artificial intelligence, machine learning and computer networks. He can be contacted at email: djnreddy@gmail.com.

Rudra Kumar Madapuri is working as professor, he received B.Tech. in computer science and engineering from Sri Venkateswara University, Tirupati, Andhra Pradesh, India. He received M.Tech. degree from Jawaharlal Nehru Technological University, Anantapur, Andhra Pradesh, India. He received Ph.D., from Jawaharlal Nehru Technological University, Anantapur, Andhra Pradesh, India. He is a life member of CSI, and he has published many research papers in national and international journals, conferences. He authored three books and also has six patents. His main research interests include software engineering, machine learning, internet of things and cloud computing. He can be contacted at email: mrudrakumar@gmail.com.