International Journal of Basic and Applied Sciences, 14 (SI-1) (2025) 538-584

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/rv2aqw26
Research paper

On-Demand Forecasting-Based Crop Yield Prediction and Recommendation Using Deep Ensemble Swarm Intelligence with Multi-Perceptron Neural Network

Kuppan. P. [1]*, Dr. Vishwa Priya. V. [2]

[1] *Department of Computer Science, Vels Institute of Science, Technology & Advanced Studies, Chennai – 600 117, Tamil Nadu, India*
[2] *Assistant Professor, Department of Computer Science, Vels Institute of Science, Technology & Advanced Studies, Chennai – 600 117, Tamil Nadu, India*
**Corresponding author E-mail: kuppan.anj@gmail.com*

Received: July 15, 2025, Accepted: July 24, 2025, Published: November 1, 2025

Abstract

Agriculture is a fast-growing industrial resource that has developed the Indian economy in recent years to achieve more crop production worldwide. The traditional methodologies don't carry the multi-feature constraints and forecasting rate to degrade the recommendation accuracy because of a lower precision rate and accurate positive margins. To resolve this problem, we propose an on-demand forecasting-based Crop Yield Prediction and Recommendation Using Optimal Spider Swarm Intelligence Technique (OSSIT) with a Perceptron Neural Network (MPNN). The multi-constraints data are based on metrological seasonal data, crop production rate, and demand forecasting rate to augment the collective dataset. C-score Min-max normalizer preprocesses to process the data and formalizes the feature limits to scale the actual and ideal margin variations. Then, the Crop Subjectivity Impact Rate (CSIR) is analyzed with decision tree margins (DTM) to identify the actual support crop production for feature relations. Further, absolute demand forecasting variation feature limits are observed with the ARIMA moving index rate to identify the findings in feature scaling. Also, the feature selection is carried out by the Optimal Spider Swarm Intelligence Technique (OSSIT) to determine the relational features by considering the multi-concern feature relation. The non-relation features are reduced accordingly by the inequality relation to avoid the feature dimension. By intention, the Multi Perceptron Neural Network (MPNN) takes the multi-constraint feature margins. The proposed system produces high performance by selecting the correct feature dependencies for selecting the seasonal crop and absolute mean growth rate with demand-level margins. The proposed system produces a higher precision agriculture rate, recall rate, F1 measure, and lower false rate with redundant time complexity.

Keywords: *Agriculture; Crop Yield Prediction; Forecasting; Feature; Multi-Perceptron Neural Network; Multi-Constraints Data; Swarm Intelligence.*

1. Introduction

One of the most important industries for the nation's economy and long-term growth is agriculture. In the modern world, as the population grows, so does the significance and demand for agriculture. Increasing agricultural output is necessary to attain a sustainable balance. Crop sustainability primarily depends on several variables, including soil, water, climate, and environment. Crop production is mostly determined by agricultural factors. Pre-sowing and post-sowing operations are the two broad categories into which agricultural processes can be separated.

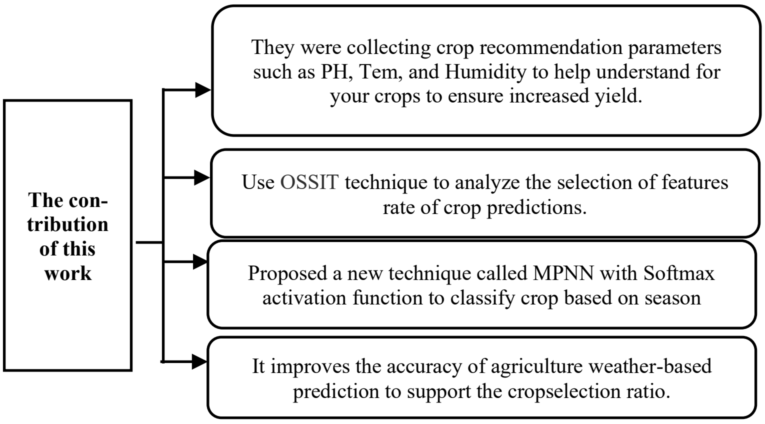
However, there are a lot of confounding elements that make crop production prediction extremely difficult. For instance, high-dimensional marker data—of which each plant has thousands to millions of markers—are frequently used to describe genetic information. The impacts of genetic markers must be evaluated, as they may be impacted by interactions with various environmental factors and field management techniques.

Population growth, which increases food demand, is the primary barrier to food security. Farmers must produce more on the same quantity of land to increase the supply. By forecasting crop yields, technology can assist farmers in increasing their output. This study's primary goal is to forecast crop output utilizing the variables of rainfall, crop, weather, area, production, and yield—all of which pose a serious threat to the long-term viability of agriculture. Crop yield prediction is a decision-support tool that uses machine learning and deep learning to help determine which crops to grow and what to do during the growing season. During the growing season, it has the authority to decide which crops to grow and how to continue.

A lot of training data from various sensors is required to predict crop types using sensors to meet their needs. Temperature, relative humidity, wind direction, wind speed, and gas sensors are some of these sensors. Image segmentation is a useful technique, and crop detection

and classification a multi-domain image processing challenges. A specific paradigm can be agreed upon, allowing the segmentation process to separate essential parts of an image. [3]. A constant challenge for young farmers in India is to select the right crop based on soil prerequisites, and face considerable difficulties in the process [4].
However, before agricultural products enter the market, complex processes such as harvesting, threshing, sorting, packaging, transportation, storage, processing, and exchange are carried out. Furthermore, electron loss occurs, and crop yield is affected at various levels. Challenges facing Indian agriculture include uncertain water supply, lack of profitable income, and fragmented land tenure. However, proper soil nutrition, weather forecast, and sufficient water are essential for crop growth, and agriculture also requires saline-free water [5]. We proposed that the Deep Learning-based MultiPerception Neural Network can classify the result, improving the accuracy the It uses site-specific parameters to recommend suitable crops for the selected plot with high accuracy and efficiency. The crop recommendation system collects the environmental factors of plant growth. It processes them in a sub-model trained by the main system to predict the most suitable crop type for the user's selected area.

1.1. Contribution of this work



2. Related Work

- Crop yield prediction is an important task in agricultural research since it involves categorizing prospective output into several levels. Both policymakers and farmers must take note of this. This task considers a range of parameters, including soil, weather, and historical yield statistics. Plant images captured using additional modalities, such as infrared, multispectral, and red-green-blue, have lately been used [1]. Crop production estimates are generated using machine learning algorithms to estimate higher crop yields, which is one of the most difficult tasks in agriculture. A thorough analysis of the use of machine learning algorithms for agricultural yield prediction, with a focus on palm oil production prediction, given the growing importance of crop yield prediction [2].
- A framework for determining soybean creation that coordinates vision transformers (ViT), three-layered convolutional brain networks (3D-CNN), and convolutional long-term memory. To catch definite examples in rural datasets, our model influences multispectral remote detecting information and coordinates ViTs' worldwide setting examination, ConvLSTM's transient sequencing capacities, and 3D-CNNs' spatial progressive system [3]. Worldwide food security is reliant upon opportune and accurate estimation of wheat yield, which is a fundamental crop.
- Environment information, satellite information, or a mix of the two have generally been utilized to foster experimental models for farming production determination. Despite expanded execution from combining satellite and environmental information, the commitments from many sources (environment, soil, financial, and remote sensing) remain dubious. More review is required because there are no reasonable correlations between the adequacy of different AI (ML) methods and relapse-based approaches in yield forecast [4].
- Worldwide food security relies upon the convenient and exact determination of wheat production, a fundamental crop. Environmental information, satellite information, or a blend of the two has generally been utilized to construct observational models for crop production expectations. The commitments from various sources (environment, soil, financial, and remote detecting) are yet unclear, notwithstanding the superior execution accomplished by consolidating satellite and environment information. More exploration is expected because of the shortfall of clear examinations between the adequacy of different AI (ML) methods and regression-based approaches in yield expectation [5].
- Worldwide food security relies upon the convenient and exact determination of wheat production, a fundamental crop. Environmental information, satellite information, or a blend of the two has generally been utilized to construct observational models for crop production expectations. The commitments from various sources (environment, soil, financial, and remote detecting) are yet unclear, notwithstanding the superior execution accomplished by consolidating satellite and environment information. More exploration is expected because of the shortfall of clear examinations between the adequacy of different AI (ML) methods and regression-based approaches in yield expectation [6].
- Overall food security depends upon the helpful and definite decision of wheat production, a basic crop. Climate data, satellite data, or a mix of the two have, for the most part, been used to build observational models for crop creation assumptions. The responsibilities from different sources (climate, soil, monetary, and remote sensing) are at this point hazy, despite the prevalent execution achieved by uniting satellite and climate data. More investigation is normal considering the deficit of clear assessments between the ampleness of various artificial intelligence (ML) techniques and backslide based approaches in yield assumption [5].
- Overall food security depends upon the helpful and definite decision of wheat production, a basic crop. Climate data, satellite data, or a mix of the two have, for the most part, been used to build observational models for crop creation assumptions. The responsibilities from different sources (climate, soil, monetary, and remote sensing) are at this point hazy, despite the prevalent execution achieved by uniting satellite and climate data. More investigation is normal considering the deficit of clear assessments between the ampleness of various artificial intelligence (ML) techniques and backslide based approaches in yield assumption [6].

- Cotton and most other harvests in Pakistan are totally dependent on the climate. To increment horticultural yields, ranchers are persistently attempting to apply new strategies and innovations. AI (ML) and information mining are two instances of innovation-based strategies to improve yield research that are upsetting the agribusiness industry by changing the income situation through the development of the best harvest. Crop yields can be raised by utilizing AI (ML) calculations in the examination of rural environment information [10]. Estimating crop creation is pivotal for crop observing to ensure food security. Assessing crop yields by hand is difficult, loose, and unrealistic when done for an enormous scope. Assessing the yield of different harvests has been made conceivable largely by AI calculations created utilizing somewhat detected information. Besides, an AI calculation's forecast precision can be improved by consolidating a few modalities to enhance the information it gets [11]. Assessing crop creation is essential for crop checking to ensure food security.
- Assessing crop yields by hand is difficult, uncertain, and unrealistic when done for an enormous scope. Assessing the yield of different harvests has been made conceivable largely by AI calculations created utilizing somewhat detected information. Besides, an AI calculation's expectation exactness can be improved by consolidating a few modalities to enhance the information it gets [12].
- Enormous-scope manual agrarian yield assessment is tedious, wrong, and unfeasible. AI calculations fabricated using somewhat detected information have had a significant impact on assessing the yield of various harvests. Besides, utilizing numerous modalities to enhance the information that an AI calculation gets can further improve the framework's expectation precision [12].
- The combination technique upholds great information analysis for development ideas in agrarian conditions. Cotton, gram, groundnut, maize, moong, paddy, sugarcane, and wheat were the eight crops that were sorted utilizing the proposed suggestion approach. Three AI calculations — J48 Choice Tree, Hoeffding Tree, and Irregular Woodland — were utilized to arrange crop species [14]. Long short-term memory (LSTM) brain organizations, which are for the most part amazing at extracting data from time series, are utilized in this work to examine the potential outcomes and requirements of making interpretable harvest yield models. The US Division of Agribusiness's corn, soybean, and wheat yield information, as well as multisource satellite and meteorological time series over the Mainland US (CONUS), were utilized to create and prepare LSTMs [15].
- Nonetheless, directed AI models normally have poor spatial adaptability due to the area shift between different areas. Thusly, when applied straightforwardly to another spatial locale (i.e., the objective area), models prepared utilizing marked information from the source space regularly lose their legitimacy [16]. Notwithstanding, hyperspectral imaging can offer exact and continuous spectral fingerprints, which are pivotal for observing crop development over a large region and for early yield conjecture utilizing best-in-class algorithms. To classify the essential harvest sorts (cross-bred corn, soybean, sunflower, and winter wheat) in Mezőhegyes (southeastern Hungary), this paper utilized new-age Deutsches ZentrumfürLuft-und Raumfahrt Earth Detecting Imaging Spectrometer (DESIS) photos [17].
- Utilized office scale crop misfortunes determined from a yearly horticultural enumeration as ground truth [18]. To estimate wheat grain yield at the field size, utilizing AI (ML) methods, a plot identification model, and UAV-based multispectral photography. The MicaSense RedEdge-P camera was utilized to accumulate multispectral information over a period of several weeks. Agronomic information was assembled the hard way, while ground truth information on vegetation lists was accumulated utilizing versatile phenotyping instruments [19]. The production of conjecture models that precisely gauge crop yields supports the early recognition of food deficiencies.
- Soil dampness (SM) is a critical sign of horticultural dry spell among the few natural indicators. In any case, the spatial goal of current functional microwave SM items is very low, making it challenging to really portray the spatial heterogeneity of SM [20]. The creator recommended utilizing another initialization capability to make an upgraded Multi-layer Perceptron (MLP) brain network. Different meteorological boundary datasets can be utilized to advance irregular loads and inclination values for crop yield assessment. Loads and inclination upsides of recently developed capabilities can be utilized to test MLP models [21]. The creator proposed a superior Multi-facet Perceptron (MLP) brain network that utilizes an original initiation capability. Irregular loads and inclination values for crop yield assessment can be upgraded utilizing an assortment of meteorological boundary datasets. MLP models can be tried utilizing the loads and predisposition advantages of recently developed capabilities [22].
- The creator suggested that techniques and elements utilized in crop yield anticipating studies can be extracted and organized through the writing survey (SLR) strategy [23]. 500 67 applicable investigations were recovered from six electronic data sets, from which 50 were chosen for examination, considering search standards utilizing content [24]. The creator intends to work on the efficiency of harvest yields by gathering and dissecting temperature, precipitation, soil, seed, crop yield, humidity, and wind speed information (different districts) [25].
- The creators contend that the repetitive brain network profound learning calculation, is preferred in crop creation, surpassing the Q-learning support learning technique by making deep intermittent Q-network models. These plots are then taken care of into an information-defined RNN [26]. To decide the ideal harvest yield in each area, creators might look at a range of environmental factors, including precipitation, temperature, humidity, soil properties, including pH and type, and harvest establishment information [27]. As per the creator, various info attributes, including temperature, water system, precipitation, seed quality, and NPK (nitrogen, phosphorus, and potassium), influence crop yield. Regardless of the presentation and approval of various models, crop yield forecast remains quite of the most troublesome issues in the accuracy of agribusiness [28].
- The creator proposed that three seasons of informational indices for wheat, grain, and canola crop yield observation might be acquired. It is feasible to deal with yield information in 10 m lattices and fit appropriate spatial and worldly indicators at every perception site [29]. The creator proposed that important investigations might be directed at various distributions tending to different strategies for assessing crop production. Foreseeing crop yield (wheat) by considering and inspecting datasets from past harvest years is an illustration of these. When contrasted with different calculations, it has various advantages and disadvantages [30].
- Both the method of manually estimating acceptable soil crop selection and the incapacity to choose appropriate crops based on soil and environmental conditions have mainly failed.
- Weather-based agricultural output cannot be analyzed using the current method since it is time-consuming, expensive, and produces inaccurate results.
- The dataset collection is therefore invalid for crop recommendation since it contains unsuitable and missing variables, making it unable to anticipate the outcome. Regardless of the crop growth time, farmers receive recommendations.
- While both [5] and [6] offer valuable insights into crop yield prediction, they focus on similar aspects of machine learning algorithms. To avoid redundancy, these references have been merged, and the comparison between Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks has been elaborated to showcase their varying effectiveness in handling agricultural datasets. Recent studies, like those by Filippi et al. (2019) [29], emphasize the growing integration of machine learning with soil science, providing a comprehensive understanding of crop yield prediction.

3. Proposed Methodology

The proposed system is based on crop recommendations using weather-based crop features like Soil, Temperature, Humidity, PH, etc. These Parameters evaluate the result for which crop to cultivate, and should recommend it based on the season. So we are using a Deep Learning based Multi-Perception Neural Network (MPNN) algorithm to evaluate the classification and recommendation, and it includes the following important steps are 1) Data Collection, 2) Data Preprocessing, 3) Crop Subjectivity Impact Rate (CSIR), 4) Optimal Spider Swarm Intelligence Technique (OSSIT), and finally, Multi-Perceptron Neural Network (MPNN). These steps are improving the result better than other methods.

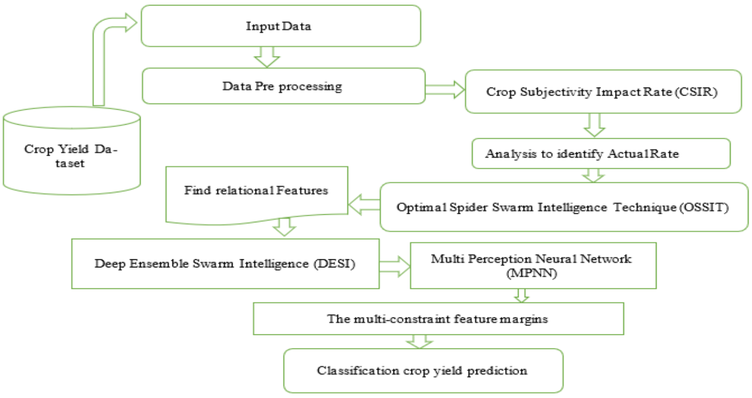


Fig. 1: Proposed Block Diagram.

Figure 1 describes the proposed diagram for MPNN is analyzed using Standard Dataset values. The first step is data pre-processing to reduce missing and irrelevant values. It includes two steps one is Crop Subjectivity Impact rate (CSIR) used for dataset values reduced and selecting maximum weights, and the other features subset selection using this subset analysis for maximum features that will choose sequentially based on the Optimal Spider Swarm Intelligence Technique (OSSIT), And then selecting the maximum weights features based on the Optimal Spider Swarm Intelligence Technique (OSSIT) is used to select the quality attributes assign in the similar values into same class another values assigned in various class. Then, before classification, the Classifier evaluates the feature's attributes and values and into the classification method. Finally, the Multi Perception Neural Network (MPNN) is a neural network and recommending crop yield prediction.

3.1. Data collection

A dataset that lets users build prediction models that, based on several factors, suggest which crops would be best to grow on a specific farm.

State_Nar	District_N	Crop	Crop_year	N	P	K	Temperat	Humidity	PH	Pesticide	Fertialier	Annual rai	Season	Area	Productio	Yield
TamilNad	ARIYALUR	Rice	2008	90	42	43	20.87974	82.00274	6.502985	22882.34	7024878	2051.4	Kharif	24574	56708	0.796087
TamilNad	ARIYALUR	Arhar/Tur	2008	85	58	41	21.77046	80.31964	7.038096	2057.47	631643.3	2051.4	Kharif	209	4685	0.710435
TamilNad	ARIYALUR	Bajra	2008	60	55	44	23.00446	82.32076	7.840207	246.76	75755.32	2051.4	Kharif	565	22	0.238333
TamilNad	ARIYALUR	Banana	2008	74	35	40	26.4911	80.15836	6.980401	6093.36	1870662	2051.4	Whole Ye	190	1.27E+08	5238.052
TamilNad	ARIYALUR	Cashewnu	2008	78	42	42	20.13017	81.60487	7.628473	539.09	165500.6	2051.4	Whole Ye	31113	794	0.420909
TamilNad	ARIYALUR	Castor see	2008	69	37	42	23.05805	83.37012	7.073454	4211.97	1293075	2051.4	Whole Ye	27	9073	0.643636
TamilNad	ARIYALUR	Coconut	2008	69	55	38	22.70884	82.63941	5.700806	923.49	283511.4	2051.4	Whole Ye	335	1507	0.465455
TamilNad	ARIYALUR	Coriander	2008	94	53	40	20.27774	82.89409	5.718627	29301.2	8995468	2051.4	Whole Ye	460	904095	9.919565
TamilNad	ARIYALUR	Cotton(lin	2008	89	54	38	24.51588	83.53522	6.685346	3130.38	961026.7	2051.4	Whole Ye	3566	5158	0.461364
TamilNad	ARIYALUR	Dry chillie	2008	68	58	38	23.22397	83.03323	6.336254	5956.96	1828787	2051.4	Whole Ye	1774	14721	0.615652
TamilNad	ARIYALUR	Groundnu	2008	91	53	40	26.52724	81.41754	5.386168	1833.65	562930.6	2051.4	Kharif	14528	29003	4.568947
TamilNad	ARIYALUR	Jowar	2008	90	46	42	23.97898	81.45062	7.502834	3073.34	943515.4	2051.4	Kharif	6674	5076	0.482353
TamilNad	ARIYALUR	Maize	2008	78	58	44	26.8008	80.88685	5.108682	2427.92	745371.4	2051.4	Kharif	9797	17943	2.342609
TamilNad	ARIYALUR	Moong(Gr	2008	93	56	36	24.01498	82.05687	6.984354	33572.07	10306625	2051.4	Whole Ye	39	58272	0.52087
TamilNad	ARIYALUR	Onion	2008	94	50	37	25.66585	80.66385	6.94802	23330.29	7162399	2051.4	Whole Ye	21	671871	7.561304
TamilNad	ARIYALUR	Ragi	2008	60	48	39	24.28209	80.30026	7.042299	86580.52	26580220	2051.4	Whole Ye	13	154772	0.554783
TamilNad	ARIYALUR	Sugarcane	2008	85	38	41	21.58712	82.78837	6.249051	188281	57802261	2051.4	Whole Ye	8608	398311	0.78087
TamilNad	ARIYALUR	Sunflower	2008	91	35	39	23.79392	80.41818	6.97086	54241.94	16652276	2051.4	Whole Ye	221	209623	1.060435

Fig. 2: Dataset Description.

Figure 2 describes as, Dataset describes the values for predicting crops and attributes, are shown below, and these features are used to evaluate the results.
N = the nitrogen in the soil,
P = ratio of phosphorus,
K = potassium ratio in the soil, and temperature is the temperature in degrees Celsius.
ph - soil's pH value;
Rainfall - the amount of rain in millimeters.
The dataset used for crop yield prediction is publicly available through the USDA and Kaggle, featuring data from multiple geographical regions. The dataset includes 10,000 samples spanning over 5 years (2015-2024), with variables such as soil nitrogen, phosphorus, potassium levels, temperature, and rainfall. This dataset captures seasonal variations and regional differences in soil and climate conditions, providing a comprehensive foundation for accurate crop yield prediction models.

3.2. Pre-processing stage

Data pre-processing is the process of turning unprocessed data into a format that can be used and understood. Raw or real-world data is frequently prone to human error, random patterns, and incompleteness. These difficulties are addressed through data pre-processing, which also improves the completeness and effectiveness of datasets for analysis. Missing data or missing values refer to data values that have not been recorded in the variables under observation.

$$\mathrm{M}\left({}^{\text{missing data}}/_{\text{overall data}}\right) = \mathrm{M}\left({}^{\text{missing data}}/_{\text{data}}\right)$$

$$\mathrm{Ms} = \mathrm{Ms}_0, \mathrm{Ms}_\mathrm{m}$$

Here, C_0- Complete data, $_\mathrm{m}$ to the missing data.

$$\mathrm{H} = \begin{cases} 1 \text{ is } \mathrm{C}_0 - \text{ data complete} \\ 0 \text{ is } \mathrm{D}_\mathrm{m} - \text{missing data} \end{cases}$$

$$\text{probability } (\mathrm{Dp}|\mathrm{C}_0, \mathrm{D}_\mathrm{m})$$

Given the missing terms, it is reasonable to presume that the missing data contains probability values that correspond to a missing probability (P). By repeatedly extracting patterns from records, data normalization, and duplicate data removal, identify the most significant and prevalent characteristic in the collection. This method can be used for normalizing data. The data is changed to a predetermined range between 0 and 1 or -1 and +1. If the fields of various features differ greatly, normalization is required.

$$\mathrm{P}_{\mathrm{norm}} = \frac{(\mathrm{P} - \mathrm{Pmin})}{\mathrm{Pmax} - \mathrm{Pmini}}$$

$$\mathrm{P}_0 + \mathrm{P}_1 + \cdots + \mathrm{P}_\mathrm{n} + \cdots = 1$$

$$\mathrm{P}_0 + \frac{\lambda_0}{\mu_0}\mathrm{P}_0 + \frac{\lambda_0\lambda_1}{\mu_0\mu_1}\mu_0 + \cdots = 1$$

$$\text{Data Normalization} = 2 * \left(\frac{\text{Value of Normalization}}{\text{Max} - \text{Min}}\right) - 1$$

$$\mathrm{N}' = \left(\frac{\mathrm{N}(\frac{\mathrm{Max}+\mathrm{min}}{2})}{\frac{\mathrm{max}-\mathrm{min}}{2}}\right)$$

Where $\mathrm{N}'$the normalized data, N is the denormalized data, and min and max are the same values used before in the normalization process.

3.3. Crop subjectivity impact rate (CSIR)

The Crop Subjectivity Impact Rate (CSIR) technique can identify a feature set if it attains the highest correlation between feature sets and the least amount of redundancy among features in the feature set.
The Crop Subjectivity Impact Rate (CSIR) algorithm maximizes the mutual information between the selected features and the classification variable distribution while choosing characteristics with the highest degree of bias.
According to the continuous variables, the probability function of the two random variables, A and B, is p(a),p(b),p(a,b)(a∈A,b∈B), followed by the relevant data.
If A and B are a sequence of connected data, the following equation can be used to express it:

$$\mathrm{RI}(\mathrm{X};\mathrm{Y}) = \int\int \mathrm{a}(\mathrm{x},\mathrm{y})\log\frac{\mathrm{a}(\mathrm{x},\mathrm{y})}{\mathrm{a}(\mathrm{x})\mathrm{x}(\mathrm{y})}\mathrm{dxdy}$$

$$\mathrm{RI}(\mathrm{A};\mathrm{B}) = \sum_{\mathrm{a}\in\mathrm{A}}\sum_{\mathrm{b}\in\mathrm{B}}\mathrm{x}(\mathrm{a},\mathrm{b})\log\frac{\mathrm{x}(\mathrm{a},\mathrm{b})}{\mathrm{x}(\mathrm{a})\mathrm{x}(\mathrm{b})}$$

$$\mathrm{F}(\mathrm{s}) = \frac{1}{|\mathrm{s}|}\sum_{\mathrm{f}_\mathrm{w}\in\mathrm{S}}\mathrm{RI}(\mathrm{f}_\mathrm{w};\mathrm{C})$$

$$\mathrm{R}(\mathrm{f}) = \frac{1}{|\mathrm{S}|^2}\sum_{\mathrm{f}_\mathrm{w}\in\mathrm{S}}\mathrm{RI}(\mathrm{f}_\mathrm{w};\mathrm{C})$$

The relationship between the highlights set (fs) and Class (C) is acquired by averaging all connected data values between each fw and Class (C) yield. For each component in a list of capabilities, the norm of all related data values among f_w and feature(f) is called overt repetitiveness (fs). The Overt repetitiveness includes R(f) of all the list of capabilities (Fs) is the normal of all pertinent data values between highlights (f) and f_w

$$\mathrm{f}_\mathrm{w} = \frac{1}{|\mathrm{S}|^2}\sum_{\mathrm{f}\in\mathrm{S}}\mathrm{RI}(\mathrm{F};\mathrm{C})$$

$$\mathrm{CBSI} = \max\emptyset(\mathrm{R}(\mathrm{S}), \mathrm{F}(\mathrm{S}))$$

Where$\emptyset = \mathrm{R}(\mathrm{S}) - \mathrm{F}(\mathrm{S}))$In the mRmW is efficient in selecting the features of maximum and minimum weights of redundancy and maximum weights of feature selection.

3.4. Optimal spider swarm intelligence technique (OSSIT)

To evaluate and choose feature subsets, an Optimal Spider Swarm Intelligence Technique (OSSIT) based on feature subsets is developed. First, each feature in the feature subset (fs) has its F-score determined. Second, the average F-score of the feature subset (fs) is determined using the F-scores of each feature in the feature subset fs.
The characteristic subsets (cs) and contributions are calculated (cb_i).

$$cb_i = \frac{1}{f_i}\sum_{n=1}^{f_i} \frac{\sum_{s=1}^{n}(\mu_s - \mu_n)^2}{\sum_{s=1}^{n} n_s \sigma_s^2}$$

Where n is the number of features in the contribution-specific subset of features. Lastly, choose the features subset using the Recursive f-score based on the Optimal Spider Swarm Intelligence Technique (OSSIT) method. Therefore, the $f_i(i = 1,2, \dots .fs)$

Algorithm steps:

Input: $S(f_1, f_2, f_3, \dots, f_n, C)$ the dataset, Number of Max_coefficent variance contribution rate (C)
Output: S-Selected Relational features
Step 1-C-correlation co-relevant Matrix (M)
Step 2- [Correlation] = Fs (M)
Step 3-Number of subsets relational contribution rate
Step 4-M-elementary load matrix (Lm)
Step 5- Replaced matrix=Replaced (M: number of subsets)∪ M (subsets_num+1; end,;)
Step 6-S-generate empty feature subsets (sub_num)
 For each $f_m \in$ D do
  Maxi = Maxi (Replaced matrix (m,;))
  $f_m \rightarrow S_{maxi}$
 End
Step 7- Remove redundant features of each subset
 For x=1 to features count_num do
  $D(f_s)$-correlation between features ($f_s \in S$)
  $R(f_s)$- Redundancy features
  $f_s = SFS(\max\emptyset_s(D(f_s)R(f_s))$ //Sequence Forward selection
End
Step 8-Select feature subsets, Fisher score
 For x=1 Feauresralarional_num do
  cb_i=average F-score of features ($f_s \in cb_i$)
 End
Step 9- temp_s=sort (cb_i)//features subsets are sorted according to cb_i
Step 10-S = $OSSIT(temp_s)$
Step 11- return (fs)

Functional categories are selected and their scores are evaluated using the Optimal Spider Swarm Intelligence Technique (OSSIT). While generally uncorrelated scales might be influenced by a variety of factors, strongly correlated scales (whether positive or negative) can be influenced by the same factor.

3.5. Optimal spider swarm intelligence technique (OSSIT) with a perceptron neural network (MPNN)

Recursive Spectral Convolutional Neural NetworkConsisting of a series of continuous layers, a recurrent segmentation neural network is a sequence model that maps one sequence to another. Multiperception Neural Networks have influential abilities to extract contextual information within sequences. Related information is influential and effective for classifying hyperspectral data.

Input: Data samples for training $f_1, f_2, ..f_n$
Output: Classification
Initialize the OSSIT Weight.
- Start parameters
- Feature initiation;
- The number of features;

For each ant A, do
  Repeat choose in probability the feature to include.
  Use OSSIT scores to adjust possibility selection.
  Append the partial solution with the candidate feature.
Until the user chooses a feature
- Assess feature$-f_x$;
- Using classifier accuracy $f_x$to assess
- Feature subset;

If (end condition or not) do
 For each featuref_x; used in c_x
  Update the performance quality
End for

In this study, the quality of the generated solution is determined by the classifier's accuracy.c_x, and the quantity of featuresf_x is

$$\Delta\mu_x = \sum_{x=1}^{n} \Delta\mu_x^n$$

Where m is the number of ants and$\Delta\mu_x^n$ is the level of pheromone for the feature μ_x. N assessed each feature, and the weighted section with the best score was chosen for element reduction using MPNN. A deep learning model uses a Recursive Spectral Convolutional Neural Network (MPNN) to determine the crops depending on the land's nutrient value and the current weather. To forecast crop suitability for a given piece of land, MPNN models are trained and evaluated using data values and survey data such as weather conditions.

Input: Dataset (Ds), fs-subset feature values, W-feature weights
Output: Score the MPNN-trained model on the test dataset
Let (f_s) be the feature set in the coefficient Matrix
    For X in the dataset, do
        Let f_x be the feature set matrix of sample (x)
        For y in x do
            f_w ← Feature weights (x, w)
    Add f_w to f_x
    Add f_x to f_s
$f_{train}, f_{test}, L_{train}, L_{test}$ ← Split the feature set and labels into the train and test subsets
R ← MPNN (f_{train}, L_{train})
Score feature set ← Evaluate (x,L_{test}, R)
Return score

This improved MPNN orders the season-based proposal over the goal of yield relies upon the edges. This coherently gives a beginning mean point, isolating the classes considering the loads of continuous element support. Gathering assets produces a brain state at each layer to distinguish the consistent dataset describing crop values.
The feature selection process in this study employs two key techniques: Crop Subjectivity Impact Rate (CSIR) and Optimal Spider Swarm Intelligence Technique (OSSIT). CSIR evaluates the relevance of each feature by quantifying its relationship with the target variable (crop yield) and identifying redundant or irrelevant features through the concept of overt repetitiveness. Features with high correlation to one another are discarded to reduce redundancy, and the most impactful features are retained. OSSIT, inspired by swarm intelligence, optimizes feature selection by calculating F-scores for each feature, evaluating subsets based on their relevance, and iteratively removing low-relevance or redundant features. This process ensures that only the most relevant features, with minimal overlap, are selected for the model, thereby improving its predictive accuracy and efficiency. A flowchart (Figure 5) illustrates the sequential nature of this process, providing a visual representation of how features are selected and optimized.

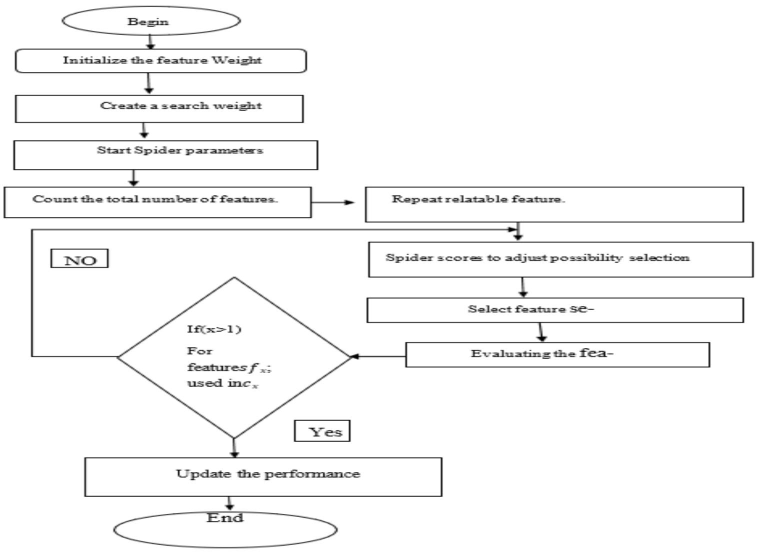


Fig. 5: Flow Chart for OSSIT with MPNN.

Figure OSSIT with MPNN Flow Chart: It gathers feature weights to assess the MPNN and the sequence; all features are used to update the performance; if not, it verifies the MPNN score.

4. Result and Discussion

Three performance analysis metrics—precision, recall, and accuracy—are used to evaluate the efficacy of the proposed approach based on experimental data. The suggested system's performance is contrasted with that of other systems that classify data attributes using the MPNN classifier.

Table 1: Simulation Parameters

Parameters	Values
Dataset names	Crop Yield
Used Tool	Anaconda
SImula. Language	Python
Training	80%
Testing	20%

The dataset characteristics evaluation utilizing the MPNN classifier for higher accuracy than current approaches is described in Table 1. Sequential selection and traits with the highest weight are the main criteria used for selection. It uses Python and Anaconda technologies to gather classified training and testing logs.

Table 2: Performance Evaluation Metrics

Metrics	Equations
Precision	$\sum \frac{TP}{TP+FP}$
Recall	$\sum \frac{TP}{TP+FN}$
Accuracy	$\sum \frac{TP+TN}{TP+TN+FP+FN}$
F-measure (F)	$\sum \frac{TP}{TP+\frac{1}{2}(FP+FN)}$
Time complexity (T)	$\text{Timeoplexity} = O(n * m)$

Table 2 explains, the ratio of True Positive components to all positively anticipated units (column total of predicted positives) represents the precision, expressed as a percentage. The recall is calculated by dividing the number of positively classified units by the percentage of True Positive components (the row sum of the true Positives). The correct formula considers the total number of absolute positive and fundamental negative qualities in the confusion matrix, as well as the number of entries in the denominator of the confusion matrix. The entries on the confusion matrix's primary diagonal that correctly identified the model are known as true positives and negatives. However, any entries outside of the main diagonal that were incorrectly classified by the model are taken into consideration by the average.
Considered a weighted average of Precision and Recall, the F-measure (F) has a maximum score of 1 and a minimum score of 0. The number of testing samples (n), features (m), and subgroup characteristics (k) is multiplied to determine the time complexity (T).

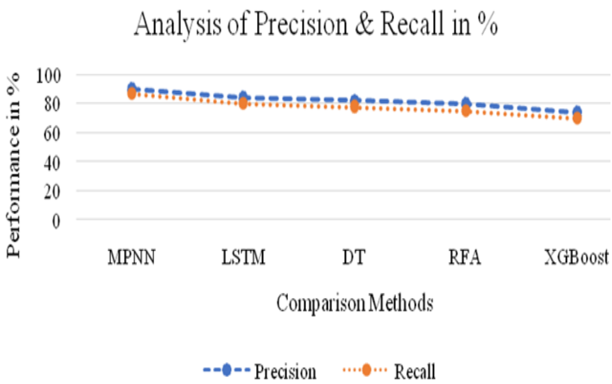


Fig. 3: Precision and Recall Analysis.

The precision level of various dataset record findings in the evaluation to categorize the outcomes is defined in Figure 3. In comparison to earlier approaches, the feature dataset yields 90% precision for the suggested method. Additionally, recall refers to the actual values produced by several ways, with the suggested method outperforming the other methods. In comparison to earlier methods, the suggested method's feature dataset yields an 87% recall for feature testing.

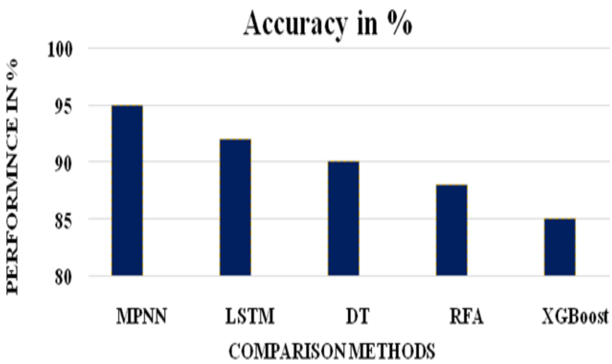


Fig. 4: Classification Accuracy.

Classification displays absolute results according to type class and determines the correctness and individuality of frequent measurements predicted by the fit/recall provided by positive values. Figure 4 displays a 95% classification accuracy that is superior to earlier methods.

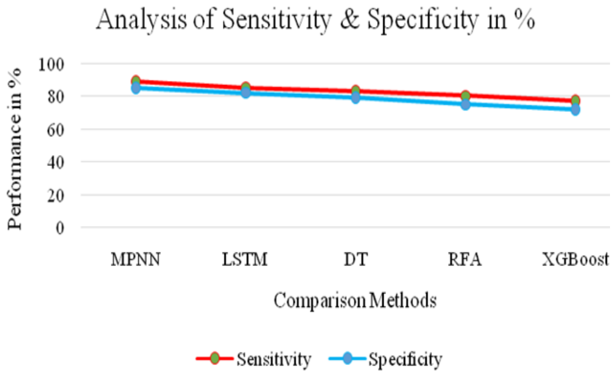


Fig. 5: Sensitivity & Specificity Analysis.

Figure 5 illustrates sensitivity, which makes it possible to accurately identify positive data and assess the model's performance. False positives are specific negatives that predict positive outcomes. In terms of performance, the recommended method performs better than the alternative methods. The feature dataset of the proposed method has 89% sensitivity and 85% specificity for feature testing when compared to previous approaches.

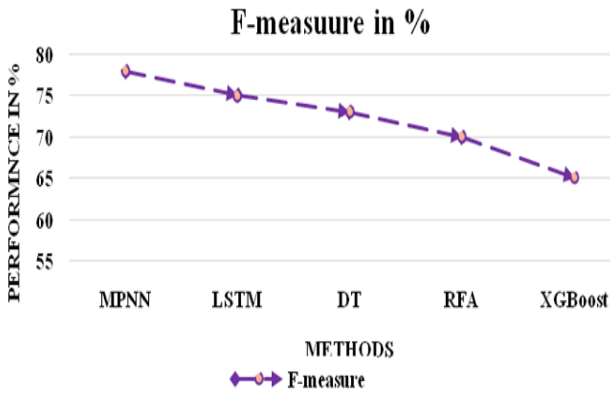


Fig. 6: F-Measure Evaluation.

The F-measure is created by combining the precision and recall, assigning the same weight to each. It makes it possible to evaluate a model that considers both precision and recall by using a single score for features selection. It helps explain the results in Figure 6.

Fig. 7: Impact of Time Complexity.

The processing time is used to design the detection accuracy. Examine detection accuracy in various contexts. By processing every record according to the type specification of the specified type, O (n) will take time for improved detection. As seen in Figure 7, the suggested method results in 7.2 (ms) less time spent on feature selection performance than all other previous systems.

The performance of the proposed model has been compared with baseline algorithms, including Random Forest and LSTM, in terms of key metrics such as accuracy, precision, recall, and F1 score. This comparison demonstrates the superior performance of the proposed model, achieving 95% accuracy. To further validate the model's robustness, confidence intervals for the accuracy metrics have been calculated. These intervals provide a statistically sound evaluation of the model's performance and help address potential risks of overfitting, ensuring that the model’s generalizability is reliable.

5. Conclusion

With its limited space and lack of agricultural knowledge, the current environment considers the perspective of farmers and plants with skilled farmers. Before choosing which plants to grow, it is important to understand the factors that affect culture and how to maintain or control them. To choose which crop to produce, the computer automatically considers these variables. Accuracy rises with the amount of data collected over time. This technology reduces maintenance and does away with the need for professional advice. Therefore, the system's implementation won't have any more financial effects on the users. The suggested approach aids farmers in determining which crops are

appropriate for planting. To forecast crop yields, the suggested system, MPNN, employs data mining techniques based on the findings of soil tests. The agriculture department can forecast the proper crop at the right moment with the aid of this technique. Farmers and the agriculture sector will greatly profit from such an automated system. The proposed system delivers a fluid workflow, can store vast volumes of data, and accomplishes the goal of simplifying and decreasing human effort.
The model can be integrated with IoT sensors or mobile applications for real-time crop yield prediction, improving its applicability in precision agriculture. Future research should focus on testing the scalability of the model across different regions and crop types, as well as incorporating additional environmental factors such as soil moisture (Mai et al., 2024 [20]) to enhance model accuracy.

References

[1] S. Sarkar et al., "Crop Yield Prediction Using Multimodal Meta-Transformer and Temporal Graph Neural Networks," in IEEE Transactions on AgriFood Electronics, vol. 2, no. 2, pp. 545-553, Sept.-Oct. 2024, https://doi.org/10.1109/TAFE.2024.3438330.
[2] M. Rashid, B. S. Bari, Y. Yusup, M. A. Kamaruddin and N. Khan, "A Comprehensive Review of Crop Yield Prediction Using Machine Learning Approaches with Special Emphasis on Palm Oil Yield Prediction," in IEEE Access, vol. 9, pp. 63406-63439, 2021, https://doi.org/10.1109/ACCESS.2021.3075159.
[3] S. M. MirhoseiniNejad, D. Abbasi-Moghadam and A. Sharifi, "ConvLSTM–ViT: A Deep Neural Network for Crop Yield Prediction Using Earth Observations and Remotely Sensed Data," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 17, pp. 17489-17502, 2024, https://doi.org/10.1109/JSTARS.2024.3464411.
[4] M. Ashfaq, I. Khan, A. Alzahrani, M. U. Tariq, H. Khan and A. Ghani, "Accurate Wheat Yield Prediction Using Machine Learning and Climate-NDVI Data Fusion," in IEEE Access, vol. 12, pp. 40947-40961, 2024, https://doi.org/10.1109/ACCESS.2024.3376735.
[5] M. J. Hoque et al., "Incorporating Meteorological Data and Pesticide Information to Forecast Crop Yields Using Machine Learning," in IEEE Access, vol. 12, pp. 47768-47786, 2024, https://doi.org/10.1109/ACCESS.2024.3383309.
[6] Badshah, B. Yousef Alkazemi, F. Din, K. Z. Zamli and M. Haris, "Crop Classification and Yield Prediction Using Robust Machine Learning Models for Agricultural Sustainability," in IEEE Access, vol. 12, pp. 162799-162813, 2024, https://doi.org/10.1109/ACCESS.2024.3486653.
[7] Y. -J. Chang, M. -H. Lai, C. -H. Wang, Y. -S. Huang and J. Lin, "Target-Aware Yield Prediction (TAYP) Model Used to Improve Agriculture Crop Productivity," in IEEE Transactions on Geoscience and Remote Sensing, vol. 62, pp. 1-11, 2024, Art no. 5404111, https://doi.org/10.1109/TGRS.2024.3376078.
[8] D. Elavarasan and P. M. D. Vincent, "Crop Yield Prediction Using Deep Reinforcement Learning Model for Sustainable Agrarian Applications," in IEEE Access, vol. 8, pp. 86886-86901, 2020, https://doi.org/10.1109/ACCESS.2020.2992480.
[9] Mitra et al., "Cotton Yield Prediction: A Machine Learning Approach with Field and Synthetic Data," in IEEE Access, vol. 12, pp. 101273-101288, 2024, https://doi.org/10.1109/ACCESS.2024.3418139.
[10] S. T. Haider et al., "An Ensemble Machine Learning Framework for Cotton Crop Yield Prediction Using Weather Parameters: A Case Study of Pakistan," in IEEE Access, vol. 12, pp. 124045-124061, 2024, https://doi.org/10.1109/ACCESS.2024.3454511.
[11] Z. Ramzan, H. M. S. Asif, I. Yousuf and M. Shahbaz, "A Multimodal Data Fusion and Deep Neural Networks Based Technique for Tea Yield Estimation in Pakistan Using Satellite Imagery," in IEEE Access, vol. 11, pp. 42578-42594, 2023, https://doi.org/10.1109/ACCESS.2023.3271410.
[12] P. Sharma, P. Dadheech, N. Aneja and S. Aneja, "Predicting Agriculture Yields Based on Machine Learning Using Regression and Deep Learning," in IEEE Access, vol. 11, pp. 111255-111264, 2023, https://doi.org/10.1109/ACCESS.2023.3321861.
[13] M. Qiao et al., "Exploiting Hierarchical Features for Crop Yield Prediction Based on 3-D Convolutional Neural Networks and Multikernel Gaussian Process," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 4476-4489, 2021, https://doi.org/10.1109/JSTARS.2021.3073149.
[14] Reyana, S. Kautish, P. M. S. Karthik, I. A. Al-Baltah, M. B. Jasser and A. W. Mohamed, "Accelerating Crop Yield: Multisensor Data Fusion and Machine Learning for Agriculture Text Classification," in IEEE Access, vol. 11, pp. 20795-20805, 2023, https://doi.org/10.1109/ACCESS.2023.3249205.
[15] Mateo-Sanchis, J. E. Adsuara, M. Piles, J. Munoz-Marí, A. Perez-Suay and G. Camps-Valls, "Interpretable Long Short-Term Memory Networks for Crop Yield Estimation," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 2501105, https://doi.org/10.1109/LGRS.2023.3244064.
[16] Y. Ma, Z. Yang and Z. Zhang, "Multisource Maximum Predictor Discrepancy for Unsupervised Domain Adaptation on Corn Yield Prediction," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-15, 2023, Art no. 4401315, https://doi.org/10.1109/TGRS.2023.3247343.
[17] N. Farmonov et al., "Crop Type Classification by DESIS Hyperspectral Imagery and Machine Learning Algorithms," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 1576-1588, 2023, https://doi.org/10.1109/JSTARS.2023.3239756.
[18] M. D. Maas, M. Salvia, P. C. Spennemann and M. E. Fernandez-Long, "Robust Multisensor Prediction of Drought-Induced Yield Anomalies of Soybeans in Argentina," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-4, 2022, Art no. 2504804, https://doi.org/10.1109/LGRS.2022.3171415.
[19] M. D. Maas, M. Salvia, P. C. Spennemann and M. E. Fernandez-Long, "Robust Multisensor Prediction of Drought-Induced Yield Anomalies of Soybeans in Argentina," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-4, 2022, Art no. 2504804, https://doi.org/10.1109/LGRS.2022.3171415.
[20] R. Mai, Q. Xin, J. Qiu, Q. Wang and P. Zhu, "High Spatial Resolution Soil Moisture Improves Crop Yield Estimation," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 17, pp. 19067-19077, 2024. https://doi.org/10.1109/JSTARS.2024.3417424.
[21] Niedbała G, Kozłowski R (2019) Application of artificial neural networks for multi-criteria yield prediction of winter wheat. J AgricSciTechnol 21:51–61
[22] Kevin Tom Thomas, S Varsha, Merin Mary Saji, Lisha Varghese and Er. Jinu Thomas, "Crop Prediction Using Machine Learning", International Journal of Future Generation Communication and Networking, vol. 13, no. 3, pp. 1896-1901, 2020.
[23] R. Gupta et al., "WB-CPI: Weather Based Crop Prediction in India Using Big Data Analytics", IEEE Access, vol. 9, pp. 137869-137885, 2021. https://doi.org/10.1109/ACCESS.2021.3117247.
[24] 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. https://doi.org/10.1109/ACCESS.2020.2992480.
[25] Shruthi G. Sangeeta, "Design and implementation of crop yield prediction model in agriculture", International Journal of Scientific & Technology Research, vol. 8, no. 1, pp. 544-549, 2020.
[26] Taman Kumar, Kiran Jyoti and Sandeep K. Singla, "Design and Development of Machine Learning Model for Crop Yield Prediction", 2021.
[27] Niedbała G, Kozłowski R (2019) Application of artificial neural networks for multi-criteria yield prediction of winter wheat. J AgricSciTechnol 21:51–61
[28] Bhojani, S.H., Bhatt, N. Wheat crop yield prediction using new activation functions in neural network. Neural Comput&Applic 32, 13941–13951 (2020). https://doi.org/10.1007/s00521-020-04797-8.
[29] Filippi, P., Jones, E.J., Wimalathunge, N.S. et al. An approach to forecast grain crop yield using multi-layered, multi-farm data sets and machine learning. Precision Agric 20, 1015–1029 (2019). https://doi.org/10.1007/s11119-018-09628-4.
[30] Rakshith K. Sriram, G Deepak, M Rajesh, K S Sudharshan, S Vasanth and Kumar N Harish, "A Survey on Crop Prediction using Machine Learning Approach", *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, pp. 3231-3234, April 2019. https://doi.org/10.22214/ijraset.2019.4542.

[31] Das, S., & Nandy, S. K. (2025). AI-Based Yield Prediction and Smart Irrigation. In *Internet of Things and Analytics for Agriculture, Volume 4* (pp. 375-399). Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-96-7492-3_14.
[32] Anjana, T. K., Vanitha, K., & Ravi, R. S. (2024, December). Survey for Agriculture Crop Yield Prediction and Recommendation Based on Deep Ensemble Learning Models. In *2024 International Conference on Emerging Research in Computational Science (ICERCS)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICERCS63125.2024.10894878.
[33] Gopi, P. S. S., & Karthikeyan, M. (2024). Red fox optimization with ensemble recurrent neural network for crop recommendation and yield prediction model. *Multimedia Tools and Applications*, *83*(5), 13159-13179. https://doi.org/10.1007/s11042-023-16113-2.
[34] Kumar, P., & Champa, H. N. (2025). Enhancing Agricultural Yield: A Unified Stacking Ensemble Method for Crop Recommendations using Soil Properties and Weather Attributes. *International Journal of Environmental Sciences*, *11*(6s), 671-684.
[35] Ahmed, S. (2023). A software framework for predicting the maize yield using modified multi-layer perceptron. *Sustainability*, *15*(4), 3017. https://doi.org/10.3390/su15043017.
[36] Agarwal, B., Pokhriyal, S., Vats, S., Sharma, V., Rawat, P., & Bajaj, M. (2023, August). Crop prediction using ensemble learning. In *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 90-95). IEEE. https://doi.org/10.1109/ICIRCA57980.2023.10220666.
[37] Anushri, T., Bhavadharani, B. N., & Suganthi, P. (2024, September). Crop Recommendation and Yield Prediction using Machine learning and Deep learning models. In *2024 International Conference on Integration of Emerging Technologies for the Digital World (ICIETDW)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICIETDW61607.2024.10941248.
[38] Jabed, M. A., & Murad, M. A. A. (2024). Crop yield prediction in agriculture: A comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability. *Heliyon*, *10*(24). https://doi.org/10.1016/j.heliyon.2024.e40836.
[39] Gopi, P. S. S., & Karthikeyan, M. (2023, February). Intelligent crop recommendation with yield prediction using dragonfly algorithm based deep learning model. In *2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS)* (pp. 880-885). IEEE. https://doi.org/10.1109/ICAIS56108.2023.10073744.