The Original

HIERARCHICAL DISCRIMINATIVE AI FOR EFFICIENT CROP RECOMMENDATION AND ON-DEMAND FORECASTING IN PRECISION AGRICULTURE

Anjana TK[1], K.Vanitha[2]

[1]Research Scholar Department of Computer Science and Engineering KAHE Coimbatore, India, anjananikhilchand@gmail.com
[2]Associate Professor Department of Computer Science and Engineering KAHE Coimbatore, India, vanitha.krishnan@kahedu.edu.in

ABSTRACT

Agriculture plays a pivotal role in country economy growth as it is major resource of food, employment and raw materials. Crop recommendations play an essential role in precision agriculture to predict the most appropriate crops. However, numerous factors such as soil characteristics, climate conditions, make it difficult to accurately predicting suitable crop. In this paper, Polynomial REgressive Momentum Optimized Discriminative AI (PREMOD-AI) model is introduced for accurately predicting crop. During acquisition phase, data samples are collected. Hierarchical Discriminative AI model includes three kinds of layers such as input layer, three hidden layers and output layer. First, collected data are given to input layer. Subsequently, data preprocessing is performedin hidden layer 1. Dimensionality reduction process is completed at hidden layer 2 by applying Hockey-stick indexed regression to determine most relevant features. Classificationis performed using Chatterjee correlation coefficients in hidden layer 3. The fine tuning process is employed to minimize error using horse herd optimization method. Finally, accurate croppredictions are obtainedat output layer. Experimentis conducted with accuracy, precision, recall, F-score, RMSE, MCC, R^2score and crop recommendation time. The result of PREMOD-AI model demonstrates an improvement in accuracy by 95%, precision by 95%, and F-score by 96%, along with a 97% increase in recall. In addition, it reduces crop recommendation time by 76ms, RMSE by 0.053, MCC by 0.91 and R^2score by 0.92 when compared to traditional deep learning approaches.

KEYWORDS: crop recommendation, Hierarchical Discriminative AI, data preprocessing, Dimensionality reduction, Chatterjee correlation coefficient based Classification, fine tuning

1. INTRODUCTION

Precision agriculture is playing a more crucial role for ensuring both sustainabilityand competitiveness in the agricultural domain. Crop recommendation for a specific region is a major challenge due to a numerous factorsincluding rapid urbanization, climate change (e.g., temperature, humidity, rainfall) soil chemistry (e.g., nitrogen, phosphorus, potassium, pH). By considering this soil composition, climate conditions, and environmental variables, farmers make decisions about which crops are most likely to be successful in specific agricultural land. The crop recommendation process not only helps in crops selection and also improves the overall productivity and minimizes resource wastage. Machine learning and advanced AI techniques have become dominant techniques for more precise crop prediction in agricultural sector.

An Improved Deep Belief Networks (IDBN) was developed in [1] for crop recommendation based on soil features, chemical soil features, climate and crop features. The Ranger Optimization was employed for selecting optimal features. But, the designed IDBNdid notconsidering short term crops and vegetables to enhance the precision of recommendation. A Gorilla Troops Optimization–based Deep Learning model for crop recommendation and yield prediction (GTODL-CRYPM) was developed in [2] aimed to enhance the prediction accuracy while minimizing the loss. However, the model exhibits high computational complexity, particularly when processing large-scale datasets, which limits its efficiency and scalability.

A lightweight exponential-weighted ensemble framework was designed in [3] that integrate numerous base classifiers to achieve better accuracy and reduce computational overhead. However, the framework did not effectively handle missing values and noisy measurements.A novel TabNet, a deep learning architecture was developed in [4] for unified smart crop recommendation by directly extracting the important patterns from preprocessed IoT-enabled agricultural data. An efficient optimization model was not applied to enhance the smart crop recommendation by reducing the misclassifications. Machine learning classification algorithms were designed in [5] for efficient Crop Recommendation by measuring the dependencies between the data. However, the algorithm did not implement an explainable AI model to enhance performance, particularly when handling large datasets. A new ensemble learning approach called random forest (RF) and extreme gradient boosting (XGB) were developed in [6] to recommend the best crop among the multiple crops. However, deep learning and machine learning models were not developed for predicting the crop yields to estimate the benefits of selecting a particular crop. Different machine learning (ML), and deep learning (DL) models were developed in [7] with the aim of recommending the crop based on structured tabular data. However, the model did not efficiently tackle the issue of computational complexity.

Explainable Artificial Intelligence (XAI) enabled Crop Recommendation system was developed in [8] by considering the various features on crop prediction. However, the smart farming systems did not adopt the user-centric approach to explain reasoning behind AI recommendations.An ensemble machine learning framework was developed in [9] by integrating entropy-based uncertainty quantification depends on agro-climatic features. Howeverthe models limited their applicability across different agricultural contexts. A tiny machine learning (TinyML) model was introduced in [10] for crop recommendation system based on statistical analysis to identify the most relevant features. However,error handling capabilities of the classifier model were not enhanced.

A Machine Learning framework's was developed in [11] for accurate crop selection and fertilizer recommendation based on specific agricultural conditions. However, the approach did not enhance the recommendations for accurately determining the quantities of macro and micro nutrients of each crop.Bayesian optimized ensemble decision tree for crop recommendation prediction (BOEDT-CRP) model was developed in [12] with various hyperparameter optimization. However, the model did not provide valuable information on the long-term effectiveness and reliability of the model under dissimilar environmental conditions. An intelligent internet of things and machine learning model was developed in [13] for crop yield prediction system by integrating the historic and current weather parameters. However, key parameters such as soil properties, fertilizer usage were not considered. Numerous machine learning models were developed in [14] for accurately perform the crop selection and determination of required nutrients. However, advanced deep learning model was not developed to increase the accuracy and a more detailed classification of different crop types. Inter-fused Machine Learning with Advanced Stacking Ensemble method was introduced [15] to enhance the crop prediction accuracy with the consideration of agricultural, environmental, and soil conditions. However, the overall processing time of the model was not reduced.

1.1 Research Problem and Motivation

Agriculture is a significant sector in numerous countries for offering food for world population. Agricultural data mining is defined as the process of extracting, analyzing and interpreting the data points from the agricultural sector to discover the hidden patterns and actionable insights. Crop recommendation is a vital field in optimizing the agricultural practices. Conventional ML as well as DL examined for performing efficient crop recommendation in agricultural field. But, the accuracy level was not improved and crop recommendation time was not reduced. But, it failed to predict the suitable crop.Despite the significant advancements of DL models in crop recommendation, hyperparameter tuning becomes necessary for enhanced performance by employing metaheuristic algorithms. Motivated by, a novel technique called PREMOD-AI is introduced for accurate prediction of crops with higher accuracy and less time.

1.2 Research Gap

Based on the review of existing literature, it is obvious that most crop recommendation systems mainly depend on machine learning and ensemble techniques, achieving improved accuracy and performance metrics. However, scalability, computational efficiency, and real-time deployment remain insufficiently addressed across multiple models. Many studies did not incorporate deep learning architectures or hybrid AI models thereby reducing the accuracy of crop recommendation and increasing the error rate. Moreover, several approaches struggled with large datasets and limited focus on time-efficient crop recommendation in sustainable agriculture. These gaps highlight the need for an advanced, scalable, and novel deep learning based crop recommendation framework integrating diverse soil composition, climate conditions, and environmental variables.

The proposed PREMOD-AI model is needed to overcome the lesser accuracy, high error rates, and computational time than conventional ML and DL frameworks in sustainable agriculture. While existing methods struggle with scalability, large datasets, and real-time deployment, PREMOD-AI introduces a highly structured, multi-layered Hierarchical Discriminative AI architecture for multi-dimensional data such as soil composition and climate conditions. Compared to traditional models, data preprocessing reduces noise.It utilizes Hockey-stick indexed regression to reduce dimensionality. It employs Chatterjee correlation coefficients for precise classification and integrates a metaheuristic Horse Herd Optimization method to automatically fine-tune hyperparameters with minimizing the prediction errors.

1.3 Key Contribution of the Model

The major contributions of the proposed PREMOD-AI modelare summarized as follows,

• To improve crop recommendation in agriculture sector, a novel PREMOD-AI modelhas been developed by integrating the Convolutional neural network and deep belief network which integrates different processes namely preprocessing, feature selection and classification
• To decrease crop recommendation time, an enhanced Convolutional Deep Belief Network (CDBN) approach is employed, which efficiently manages data preprocessing and selects relevant features.
• The preprocessing phase eliminates outliers, thereby reducing the overall dataset size. Furthermore, the Hockey-stick indexed regression is utilized to identify and retain significant features while discarding less relevant ones, improving model efficiency and crop predictive performance.
• The PREMOD-AI modelincreases the accuracy and minimizes the prediction errors, by usingChatterjee correlation coefficients for precise crop prediction. Additionally, an stochastic adaptive moment algorithm and horse herd optimization are applied to further refine the hyperparameters by minimizing the error rate.
• Finally, comprehensive experimental evaluations across various performance metrics demonstrate that the PREMOD-AI modeloutperforms conventional deep learning methods.

1.2 Paper Organization

The rest of this paper is partitioned into the subsequent sections:
Section 2 provides a literature review, summarizing previous research of crop recommendation, highlighting their contributions, and identifying limitations. Section 3 details the PREMOD-AI model, describing the task of each part, including the design of the crop prediction and development of the deep learning model. Section 4 presents the dataset description and implementation result. To end with, Section 5 concludes the paper by summarizing the key contributions and findings.

2. RELATED WORKS

Long Short term Memory model was developed in [16] for accurate Crop recommendation and forecasting system. However, scalability of the model was not improved for precision agriculture in diverse agricultural regions. An integration of Machine Learning (ML) and Internet of Things (IoT) were developed in [17] to enhance the agricultural within a Crop Recommendation System. But it did not incorporating additional parameters relevant to crop health and soil conditions to enhance crop recommendations' accuracy. Machine learning models were developed in [18] with the aim of improving the crop recommendation and yield-prediction using soil and environmental parameters collected from various sensors deployed in agricultural fields. An integration of AI and IoT was developed in [19] to provide suitable crop recommendations for optimal resource allocation and maximize yields. However, the model did not focus on refining data collection methodologies and exploring additional parameters to improve the precision and reliability of the system. Optimization based ensembles of different methods were introduced in [20] to improve the precision and stability of crop prediction. But, the method did not incorporate deep learning based crop prediction.
Gradient Boosting algorithm was developed in [21] for accurately recommend the crops depends on nutrients and environmental parameters. However, developing explainable, reliable model was major issues to enhance agricultural productivity. Two robust machine learning architectures were developed in [22] for classification and regression for accurate crop recommendation by considering the various climate and soil attributes. However, it failed to improve the accuracy of models' while handling large dataset. Machine Learning (ML) based system was designed in [23] based on historical data on climatic circumstances, soil properties, crop yields, to provide personalized crop recommendations. However, the system did not focus on incorporating more detailed and temporal soil and climate data, as well as historical crop planting information to improve the adaptability and accuracy. A stacking ensemble model was introduced in [24] to enhance crop recommendation based on feature fusion. The model failed to incorporate more comprehensive parameters such as the specific water requirements for individual crops within deep learning architectures. An integration of ensemble learning (EL) and multiple machine learning (ML) models were developed in [25] for improved the performance of crop recommendation. However, computational complexity of the model remained limited in large-scale dataset.
Machine learning techniques were developed in [26] with the aim of accurate crop recommendation based on soil nutrient analysis under diverse environmental regions and conditions. However, the model failed to perform soil fertility analysis for fertilizer recommendations to farmers.Recursive Spectral Convolutional Neural Network (RSCN2) model was developed in [27] for sustainable crop recommendation to optimizes the resource and feature subsets. However,time efficient crop recommendation was major concern. A Machine Learning-Enabled System was developed in [28] for efficient crop recommendationsto achieve high yield. However, it failed to handle different soil parameters and optimize crop types. Transformative Crop Recommendation Model was introduced in [29] to optimize recommendations depends on environmental and agronomic factors. However, providing accurate and timely recommendation was major issues. An automatic water irrigation system was introduced in [30] for accurate crop recommendation depends on soil parameters, aiming to enhance the agricultural productivity. However, an efficient deep learning algorithm was not applied to minimize the error in crop recommendation. Bi-phasic deep learning

framework called AgroDualNet was designed in [32] for optimizing the yield quality and minimizing the agricultural losses. But, the accuracy was not enhanced. To develop farming idea, new framework termed adaptive neuro-fuzzy inference system (ANFIS), support vector machine (SVM), and Puma optimizer (ANFIS-SVM-PO) was introduced in [32]. However, it failed to minimize time. Crop recommendation system was estimated in [33] with higher accuracy. But, the error was not concentrated. Table 1 shows thecomparison analysis of state-of-the-art methods.

Table 1 Comparison Analysis Of State-Of-The-Art Methods

Reference . No	Author name	Methods	Objective	Pros	Cons
[16]	Yashashree Mahale., et al., (2024)	Long Short term Memory model	Accurate Crop recommendation and forecasting system	Increase Accuracy of Crop recommendati on	scalability of the model was not improved
[17]	Mohamed Omar Abdullahi., et al., (2024)	Machine Learning (ML) and Internet of Things (IoT)	To enhance the optimal Crop Recommendation System	Increase Accuracy, precision, recall, F1score	Did not incorporating additional parameters relevant to crop health and soil conditions to enhance accuracy
[18]	Mohamed Bouni., et al., (2024)	Machine learning models	Improve the crop recommendation and yield-prediction using soil and environmental parameters	Increase Accuracy, precision, recall, F1score	deep learning techniques was not employed for accurate predictions
[19]	Md Anisur Rahman., et al., (2024)	An integration of AI and IoT	To provide suitable crop recommendation for optimal resource allocation	Enhance predictive accuracy	Did not focus on exploring additional parameters to improve precision and reliability of system
[20]	Sasmita Subhadarsinee Choudhury., et al., (2024)	Optimization based ensembles of different methods	To improve precision and stability of crop prediction	Prediction accuracy was enhanced	Did not incorporate deep learning based crop prediction
[21]	Sourabh Shastri., et al., (2025)	Gradient Boosting algorithm	Accurately recommend crops depends on nutrients and environmental parameters	Increases the precision an recall	Developing explainable, reliable model was major issues to enhance agricultural productivity
[22]	Abid Badshah., et al., (2024)	Two robust machine learning architectures	To enhance classification and regression for accurate crop recommendation	Increases the precision an recall and minimize RMSE	Failed to improve accuracy of models' while handling large dataset
[23]	Farida Siddiqi Prity., et al., (2024)	Machine Learning (ML) based system	To provide personalized crop recommendations	Increase the accuracy of Prediction performance	system did not focus on incorporating more temporal soil and climate data
[24]	Rania A. Ahmed., et al., (2024)	stacking ensemble model	To enhance crop recommendation based on feature fusion	High classification performance	Failed to incorporate more comprehensive parameters for individual crops within deep learning architectures.
[25]	Hemalatha Gunasekaran., et al., (2024)	integration of ensemble learning (EL) and multiple machine	To enhance performance of crop recommendation	Improved accuracy and reduced error rate	computational complexity of the model was not reduced

		learning (ML) models			
[26]	Sehrish Munawar Cheema and Ivan Miguel Pires (2025)	Machine learning techniques	Accurate crop recommendation based on soil nutrient analysis	Enhanced prediction performance	Failed to perform soil fertility analysis for fertilizer recommendations
[27]	Gopinath Selvaraj., et al., (2025)	Recursive Spectral Convolutional Neural Network (RSCN2) model	sustainable crop recommendation	Less error rate	Rime efficient crop recommendation was major concern
[28]	Pedina Sasi Kiran., et al., (2024)	Recursive Spectral Convolutional Neural Network (RSCN2) model	Efficient crop recommendation and yield	Enhanced precision, recall, F1-score and accuracy	Failed to handle different soil parameters and optimize crop types
[29]	Gurpreet Singh & Sandeep Sharma (2025)	Transformative Crop Recommendation Model (TCRM)	Optimize recommendations depends on environmental and agronomic factors	Increase precision , recall, F1 score	Accurate and timely recommendation was major issues
[30]	Anitha Palakshappa., et al., (2024)	Automatic water irrigation system	Accurate crop recommendation depends on soil parameters	Maximizing accuracy	Error rate was not improved
[31]	Mohanappriya., et al.	AgroDualNet	Optimize yield quality	Reduce the error	Accuracy was not enhanced
[32]	Thandra Jithendra et al.	ANFIS-SVM-PO	Forecast optimal rainfall	Precision was enhanced	Time was lesser
[33]	S. Poornima et al.	Crop recommendation system	Select most suitable crops	Higher accuracy	Error was not concentrated

Based on the above conventional literature, the limitations are described such as high computational latency [25, 27], poor scalability [16, 22], and a failure to minimize overall prediction errors [30, 31]. Also, the deep learning models such as RSCN2 [27, 28] were utilized to offer strong evaluation metrics. But, the accuracy and time is major practical concern for appropriate crop prediction. To address this issue, the proposed PREMOD-AI model by integrating preprocessing, feature selection, and classification and optimization. Unlike previous approaches that struggle with massive datasets [22], PREMOD-AI systematically minimizes recommendation time and dataset size by preprocessing and Hockey-stick indexed regression to filter out outliers and irrelevant features and pick pertinent features. Furthermore, while contemporary models fail to adequately address error rates or explainability [21, 31], the proposed work utilizes Chatterjee correlation coefficients, stochastic adaptive moment and horse herd optimization algorithm. This dual-layer approach dynamically tunes hyperparameters and lowers prediction errors. Finally, the crop recommendation into a highly robust, scalable, and time-efficient solution are achieved with maximum accuracy.

3. PROPOSAL METHODOLOGY

Agriculture is an essential role in India's economy, sustaining millions of livings. However, traditional farming methods often resist with making accurate decisions about which crops to cultivate to gain the better food productivity. This lack of precision reduces the productivity and hence novel smart technologies is required that assists the farmers decide their right crops quickly and forecast agricultural outcomes with higher accuracy. To address this concern, a novel Hierarchical Discriminative AI model called PREMOD-AI is introduced. This Hierarchical Discriminative AI method is adapted for predicting various crops to cultivate for particular land based on sol and weather conditions.

PREMOD-AI model stands for Polynomial REgressive Momentum Optimized Discriminative Artificial Intelligence. The term **polynomial regressive**highlights the framework's ability to perform the missing data imputation with other available data. The term stochastic adaptive **Momentum** algorithm is used to reduce errors and enhance the accuracy of crop prediction with the help of Horse herd metaheuristic **optimization** based hyperparameter fine tuning. **Discriminative Artificial Intelligence** is designed to classify or predict labels for input data by learning the given input data samples. Here the hybrid Convolutional Deep Belief Network (CDBN) approach is used as Discriminative Artificial Intelligence model for accurate crop prediction to support reliable and sustainable agricultural planning.This proposed PREMOD-AI model integrates the four phases such as data preprocessing, feature selection, classification, and fine-tuning to achieve more reliable and interpretable crop prediction for efficient recommendation.

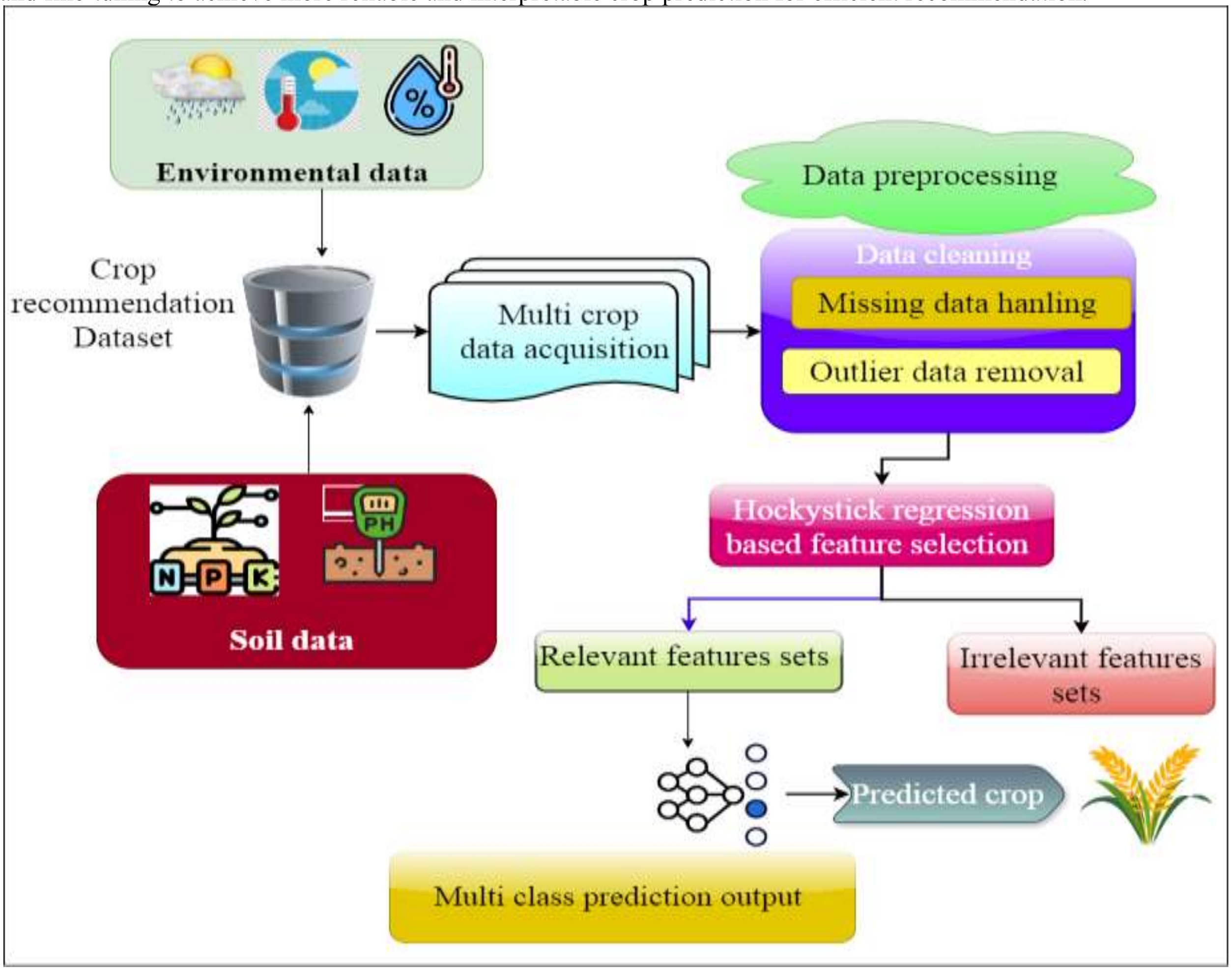


Figure 1 Architecture Of Proposed PREMOD-AI Model

Figure 1 shows the overall architecture of the proposed PREMOD-AI model for accurate crop prediction from the given environmental and soil data. The above structure includes four significant processes namely Data acquisition, pre-processing, feature selection, and classification integrated into the structure of Hierarchical Discriminative AI model to enhance the accuracy of crop recommendation. To begin with, large volumes of environmental and soil data samples are collected from the crop recommendation dataset for accurate crop prediction. The primary step of proposedPREMOD-AI model involves data pre-processing or cleaning to transform raw dataset into more suitable form through missing data handling and outlier data removal. Following the preprocessing step, the significant feature selection and removal process is carried out by employing the hockystick indexed regression function for accurate crop prediction with minimal time. To end with, the proposed PREMOD-AI model performs the crop prediction through classificationattaining higher accuracy with lesser errors. These different fundamental processes of the proposed PREMOD-AI model are explained briefly in the subsequent sections.

3.1 Data Acquisition

The first process of proposed PREMOD-AI model is the data acquisition which involves collecting the raw data samples from the dataset. In the proposed model, raw data samples are collected from the crop recommendation dataset https://www.kaggle.com/datasets/irakozekelly/crop-recommendation-dataset. This dataset consists of essential information for accurate crop recommendation and it also including the features about the soil, weather conditions, and region-specific information. It includes 18079 records and 8 features or attributes. The 8 features in this dataset includes record ID, soil factors like pH, potassium (K), phosphorus (P), and nitrogen (N) levels, as well as climate factors such as temperature, rainfall, and crop. The feature descriptions are listed in table 2 and table 3.

Table 2 Description Of Features

S. No	Features or attributes	Description
1.	**ID**	ID for that particular record
2.	N	Nitrogen content in soil
3.	P	Soil phosphorus content
4.	K	Soil Potassium content
5.	Temperature	Average temperature (in °C) impacting crop growth cycles
6.	ph	Soil pH level (affecting nutrient availability to plants)
7.	rainfall	Average rainfall amount (in mm) influencing water availability
8.	Crop	Recommended crop type based on the given soil and weather parameters **40 unique crop type** 'barley' 'sunflower' 'sweetpotato' 'rice' 'soyabean' 'jowar' 'potato', 'turmeric' 'jute' 'maize' 'moong' 'onion' 'cabbage' 'ragi' 'banana', 'rapeseed' 'wheat' 'horsegram' 'coriander' 'mango' 'tomato' 'cotton', 'brinjal' 'garlic' 'blackgram' 'cardamom' 'blackpepper' 'orange', 'bittergourd' 'papaya' 'grapes' 'ladyfinger' 'pineapple' 'bottlegourd', 'jackfruit' 'pumpkin' 'cauliflower' 'cucumber' 'drumstick' 'radish'

Table 3Number Of Data In Class Label

Crops	Number of data	Crops	Number of data	Crops	Number of data	Crops	Number of data
Rice	2245	cotton	613	blackpepper	119	cauliflower	23
Maize	1976	sweetpotato	576	mango	66	cucumber	20
Moong	1327	ragi	557	tomato	59	grapes	19
wheat	1287	horsegram	545	papaya	56	jackfruit	18
rapeseed	1118	turmeric	528	brinjal	55	drumstick	17
potato	1083	banana	507	cardamom	54	bottlegourd	13
jowar	1055	coriander	473	ladyfinger	41	radish	13
onion	1047	soyabean	442	orange	27	blackgram	11
sunflower	737	garlic	415	cabbage	26	bittergourd	11
barley	654	jute	211	pineapple	26	pumpkin	9
Total	12,529	**Total**	4867	**Total**	529	**Total**	154

3.2 Hierarchical Discriminative AI model Enabled crop prediction

After the data acquisition, proposed PREMOD-AI model performs the crop prediction which employs a Hierarchical Discriminative AI model. It is a deep learning model to achieve fast and highly accurate crop prediction with reduced computational overhead. This enhanced deep learning method is specifically developed to boost prediction performance while reducing overall computation time. By integrating the layered learning capability of Deep Belief Networks (DBNs) and strength of Convolutional Neural Networks (CNNs), the Convolutional Deep Belief Networks is employed as a Hierarchical Discriminative AI model efficiently enhance both feature selection process as well as classification outcomes. The overall architecture consists of several stacked Restricted Boltzmann Machine (RBM) layers, which helpsforaccurate learning by minimizing the training errors.

In this PREMOD-AI methodology, the CNN component efficiently extracts high-level spatial and environmental features from soil, climate, and agronomic data. Subsequently, the DBN model performs deep hierarchical learning to capture complex nonlinear relationships and temporal dependencies among extracted features. This hybrid architecture enhances crop prediction accuracy, stability, and generalization compared to separate models. Moreover, PREMOD-AI improves scalability, minimizes the prediction error rates for sustainable and time-efficient agricultural planning. The overall structural design of the CDBN is illustrated in Figure 2.

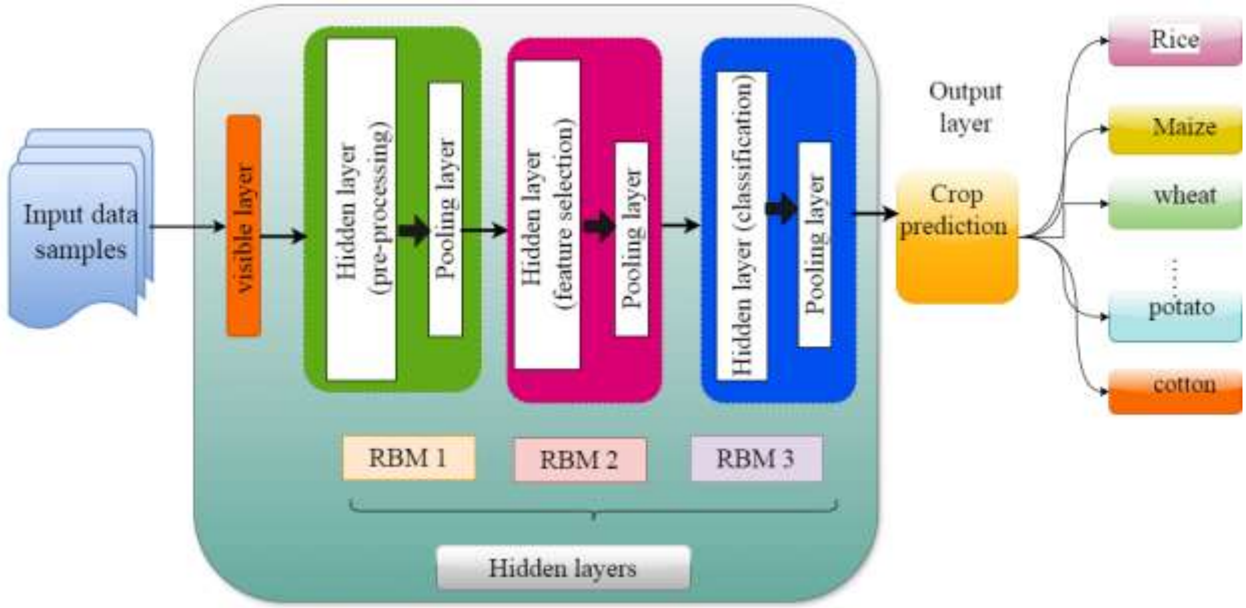


Figure 2 Schematic Structure Of A Convolutional Deep Belief Network

Figure 2 given above depicts the structure of anintegrated **Co**nvolutional Deep Belief Networkdesigned for precise crop prediction in agriculture domain. The newly developed integrated constructionincludes two main steps namely layer-wise pre-training process and fine-tuning. In the layer-wise pre-training process, each layer in network architecture is trained separately to learn a given input data samples by means of weighted computations, and the resulting processed output is transformed to the next hidden layer. During the fine-tuning stage of proposed deep learning architecture, the complete network structures perform the backpropagation to optimize the model hyperparameters, aiming to reduceprediction errors and improve overall accuracy performance in agriculture crop prediction.

The integrated deep learning network utilizes the generative stochastic artificial neural network called Restricted Boltzmann Machines (RBMs) for the layer-wise pre-training process. As shown in figure 2, there are three RBM layers are used for performing the three different tasks namely preprocessing, feature selection and classification. Each RBM includes alternating hidden and pooling layers. The small individual neurons in the hidden layer obtain the input data samples from previous visible layer or input layer. The pooling layer, followed by the hidden layer, reduces the dimensionality of the original dataset.

First, the numbers of collected data are given to the input layer or visible layer of the **Co**nvolutional Deep Belief Network. Subsequently, data preprocessing is carried out in the first hidden layer of RBM 1 to convert raw agricultural data into a clean and normalized format, enabling accurate crop prediction. In the second hidden layer of RBM 2, feature selection is performed to identify the most relevant features while eliminating redundant attributes. The third hidden layer in RBM 3 carries out the classification process to achieve higher prediction accuracy.

To begin with, the network receives input data samples, with each layer consisting of neurons that transfer the processed results from one layer to the subsequent layer. The visible layer receives the raw data samples directly, and the output from one RBM given to the input of the visible layer of the next RBM, enabling a hierarchical, layer-by-layer learning process, as shown in Figure 2.

As shown in the figure 2, the overall deep learning architecture have the training set $\{AS_i, Y_k\}$ where AS_i indicates a training agricultural data samples$AS_i = \{AS_1, AS_2, \dots, AS_n\}$' collected from the crop recommendation dataset and a label or output 'Y_k' representing its prediction output for recommending the suitable crop. The significance of a neuron activation probability play a vital role, refers to the likelihood that a neuron becomes active (produces a significant output) when given input data. It helps to determine how quickly a neuron responds across all input data samples. The activation probability of a hidden unit is computed based on a weighted sum of the inputs collected from the visible units and a bias. This weighted sum is then transferred to the sigmoid function that generates the conditional probability to verify the neurons in the hidden unit areactive or not. Therefore, the neuron activation probability in the hidden layer 'P_h' is given below,

$$P_h = \sigma\left(\sum_{i=1}^{n} AS_i \quad * w_{vh}\right) + b (1)$$

Where, P_h denotes a neuron activation probability in the hidden layer, AS_i denotes ainput data samples associated to a weight between visible and hidden layer 'w_{vh}' and included with bias 'b' helps to shift the input data samples left or right, giving it more flexibility, σ represents a sigmoid activation function. When the activation probability $P_h$becomes 1, the neuron is always active for every input sample.

3.2.1 Data Pre-processing

Data pre-processing is a second vital step of proposed PREMOD-AI model that helps to make a dataset is clean, reliable, and appropriate for accurate cropprediction. Since raw agriculture data collected from external repositories often comprises a noise, missing values, and uneven patterns. This inconsistency of the dataset leads to high error in accurate crop prediction. Therefore, the proposedmodel performs preprocessing to enhance data quality. The preprocessing steps of proposed model mainly include two major processes namely missing data handling and outlier data removal.

The initial stage of proposedPREMOD-AI model focuses on identifying and addressing missing values. A missing value is plainly unrecorded or not present for a particular variable in a dataset, appearing as blank cells. To handle these data effectively, a polynomial proximal regressive imputation method is applied to determine and replace the missing data within the input dataset. The novelty of the polynomial proximal regressive imputation is integrated within the proposed PREMOD-AI framework. This preprocessing step is efficiently hybridized within the layers of CDBN architecture to handle the missing values, and varying data ranges present in agricultural datasets. This combined approach ensures consistency across heterogeneous soil, climate, and environmental parameters. The hybrid strategy enhances feature representation and supports more stable and accurate learning in the PREMOD-AI framework thereby reducing error rates and improving overall crop recommendation performance.

Let us consider the crop recommendation dataset 'DS' and data samples 'AS' as well as features $\{f_1, f_2, \ldots, f_m\}$are arranged in the form of matrix. Therefore, the input matrix is formulated as given below,

$$IM = \begin{bmatrix} f_1 & f_2 & \cdots & f_m \\ AS_{11} & AS_{12} & \cdots & AS_{1n} \\ AS_{21} & AS_{22} & \cdots & AS_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ AS_{m1} & AS_{m2} & \cdots & DS_{mn} \end{bmatrix} \quad m = rows, n = column \quad (2)$$

Where,IM represents an input data matrix where each column indicates a number of features$f = \{f_1, f_2, \ldots, f_m\}$, each row represents a number of records or data samples or records '$AS = \{AS_1, AS_2, \ldots, AS_n\}$' respectively. First, detect where the missing values (often represented as none, or blanks are located in given input dataset. After detecting the missing values, the polynomial proximal regressive imputation process is applied for commuting the missing values with the other known data samples.

$$AS_{imputed} = \beta_0 + \beta_1 AS_1 + \beta_2 {AS_2}^2 + \cdots \beta_n {AS_n}^n + \epsilon \quad (3)$$

Where, $AS_{imputed}$ indicates a imputed output, $AS_1, AS_2, \ldots, AS_n$ denotes a number of known data samples or instances, $\beta_0, \beta_1, \beta_2, \ldots \beta_n$ indicates a regression coefficients, ϵ indicates the error term which minimizes the sum of the squared differences between the actual imputed value and predicted imputed value.

$$\epsilon = arg\,min\left(AS_{imputed}(actual) - AS_{imputed}(predicted)\right)^2 \quad (4)$$

The regression function finding the appropriate values of the coefficients that reduces i.e. $argmin$ the least absolute deviation between the actual $AS_{imputed}(actual)$'and predicted values '$AS_{imputed}(predicted)$'. In this way, regression based missing data smoothing is efficientlyperformed in the given dataset.

Followed by, an outlier detection and removal process is conducted in proposed PREMOD-AI model by means of the Extreme normalized residual test to discover outlier data points, which considerablydiverge from the rest of the other data points in the given dataset. The novelty of using an Extreme Normalized Residual Test Statistic in PREMOD-AI model for outlier detection in the hidden layer of a CDBN lies in identifying and handling outliers during internal feature learning. This approach detects extreme deviations in hidden-layer activations and identifies noisy or outliers patterns. As a result, the model achieves better accuracy and reliability in crop recommendation.

Therefore, the normalized residual test is a statistical method for efficiently detecting the outlier, and its test statistic is expressed as follows,

$$ENRT = arg\,max\left[\frac{|AS - \mu|}{v\,(AS, \mu)}\right] (5)$$

$$\mu = \frac{1}{n}\sum_{i=1}^{n} AS_i (6)$$

$$v\,(AS, \mu) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(AS_i - \mu)^2} (7)$$

Where, $ENRT$ denotes aExtreme normalized residual test statistic, μ represents a mean, v designates a deviation, n indicates a number of input data samples, $arg\,max$ denotes a argument of maximum function. Therefore, the maximum of these standardized test statisticoutcomes from the mean is said to be an outlier. In other words, the greatest absolute deviation between the data and mean is said to be an outlier. The test is continued iteratively until detecting all outliers within the given dataset. These so-called outliers are then removed from the dataset. The pooling layer followed by the hidden layer reduces the dimensionality of the originaldataset by removing the outlier samples. These preprocessed dataset results are transferred into the next RBM layer.

3.2.2 Feature Selection

The proposed PREMOD-AI model executes the feature selection process in the hidden layer of Convolutional Deep Belief Network **to** minimize the overall size of input dataset for accurate crop prediction with minimal time consumption. The main aim of this process is to preserve the more significant environmental or soil features for accurate analysis and modeling. The selected less number of significant features helps to reduce the computational performance in crop prediction. The proposed PREMOD-AI model utilizes the Hockey-stick indexed regression for accurate feature selection. The Hockey-stick regression is employed for measuring the linear relationship between input variables (features) and an output variable (target).

The proposed regression utilizes the Roger similarity statistical index function to analyze the relationship between the features and a target.This helps to project the more relevant features from high dimensional space into a low-dimensional space and also minimize time complication of the algorithm to perform accuracy of crop prediction. Let us consider the set of features in the dataset.

$A_j = \{A_1, A_2, \ldots, A_m\} where j = 1,2,3 \ldots m$ (8)

The similarity relationship between the target variable and the input soil and environment features is computed through following expression.

$SI\left(A_j,\ Y_k\right) = \frac{|A_j \cap Y_k|}{m+|A_j \,\Delta\, Y_k|}$ (9)

Where$SI\left(A_j,\ Y_k\right)$ denotes a Roger similarity index output,$A_j \cap\ Y_k$indicatesamore association between the feature 'A_j' and target output 'Y_k', $A_j \,\Delta\, Y_k$representsadissimilarityor variation between the features and target output. The similarity index function returns the outcomes between 0 and 1. The Hockey-stick regression method is a statistical approach used to distinguish between relevant and irrelevant features in a dataset based on their similarity index output. The method works by comparing the observed similarity of each feature. Features whose observed similarity exceed the corresponding Hockey-stick thresholds are considered relevant, while those lessening below the threshold are considered irrelevant. By selecting only the relevant features, this method helps to reduce dimensionality, eliminate irrelevant features and improve the performance of prediction tasks.

$A_{RE} = c_1.SI + \tau_1, if SI < HT$ (10)

$A_{SEL} = C_2.SI + \tau_2, if SI > HT$ (11)

Where,'A_{RE}' represent the removed features sets,A_{SEL} denotes a selected features sets, SI Roger similarity index, regression coefficient 'c_1', 'C_2' and regression constants 'τ_1', τ_2' with regards to Hockey-stick thresholds (HT). From the evaluation results, 'A_{SEL}' results are considered for accurate crop prediction. The selected features results are transferred into thirdRBM layer.

3.2.3 Classification

After selecting the significant features from the dataset, the final step of proposedPREMOD-AI model is to perform the classification in third RBM layer of network architecture along with the selected features. In the third RBM layer, training and testing data analysis is done by the means ofChatterjee correlation coefficient to classify the crops for accurate recommendation.

$CRC = 1 - \frac{|AS_{Tr} - AS_{Ts}|}{ST(AS)}$ (12)

$ST(AS) = \frac{|AS_{Tr} - AS_{Ts}|^2}{|AS_{Tr} - AS_{Ts}|^2 + \rho^2}$ (13)

$\rho = \frac{AS_{Tr} - AS_{Ts}}{\sqrt{(1-AS_{Tr})^2(1-AS_{Ts})^2}}$ (14)

Where, CRC denotes a Chatterjee correlation coefficients, AS_{Tr} denotes a training data samples, AS_{Ts}denotes a testing data samples, $ST(AS)$indicates wald statistic test statistics, $V(\ A_j,\ TR_k)$ represent variance between the training and testing samples. The correlation 'CRC'returns the output ranges from '0 to '1'. If the outcome of 'CRC' is more close to 1, indicates that the data samples have been accurately predicted. For each crop classification, the error '$CFE\,'$ is computed as follows

$CFE = \frac{Misclassified\ data\ samples}{Total\ number\ of\ samples}$ (17)

During the fine-tuning stage of proposed model, the hyperparameters of deep learning structures are updated with the stochastic adaptive momentum algorithm toreduce errors and enhance the accuracy of cropprediction. This algorithm helps to iteratively update the weights between the layers of deep learning model for faster and more efficient convergence.

$w_{new} = w_t - \eta \frac{Z_t}{\sqrt{Q_t}+\varepsilon}$(18)

Where, $Z_t = \gamma_1 Z_{t-1} + (1-\gamma_1)\left[\frac{\partial CFE}{\partial w_t}\right]$ (19)

$$Q_t = \gamma_2 Q_{t-1} + (1-\gamma_2)\left[\frac{\partial CFE}{\partial w_t}\right]^2 \quad (20)$$

Where, w_{new} denotes an updated weights between the layers, w_t represents a current weight, η indicates a learning rate($\eta < 1$), $\frac{\partial CFE}{\partial w_t}$' indicates a partial derivative of the classification error 'CFE' regarding current weight 'w_t',γindicates a Momentum parameter (0. 9), Z_t represents a first moment (i.e. mean), G_t indicates a second moment (i.e. variance), ε
denotes a small positive constant added to the denominator. From the updating process, numerous weights are observed based on the computed gradients. Therefore, an optimal weight is determined by applying Horse herd algorithm to enhance the accuracy by decreasing the error (i.e. false positive or false negative).

Horse herd algorithm is a current population-based metaheuristic optimization approach aimed to address more complex issues in large volume of data samples. It imitates the social dynamics and movement patterns monitored in horse herds to investigate the search space efficiently.
In this algorithm, the term weights related to the horses. Initialize the population of the horses (i.e. number of weights 'w_r') in search space.

$$w_r = w_1, w_2, w_3 \ldots . w_r \quad (21)$$

Once the population is initialized,composure the fitness based on classification error.

$$f(w_r) = arg\ min\ CFE \quad (22)$$

Where $f(w_r)$indicates a fitness of each weight value, $arg\ min$ indicates an argument of minimal function, CFEindicatesa error rate in crop classification. After the fitness calculation, the current best horse i.e. weight is chosen among the others for selecting the optimal. The current best horse is chosen as leader in their herd. Once the current best is chosen, different horse behaviors such as exploration and exploitation are discussed in the following subsections.

3.2.3.1 Exploration Phase

During the exploration stage of the optimization process, the horses (search agents) find the way in a random and diverse manner to discover promising new solution. This random exploratory movement helps to prevent early convergence and ensures that the algorithm comprehensively observes the search landscape

$$X_{t+1} = X_t + R_1 * 0.5|X_{best} - X_t| \quad (23)$$

Where,X_{t+1}denotes a updated position of the horse, $X_t$represents a current position of the horse, R_1 denotes a random number (0, 1), $0.5|X_{best} - X_t|$ represents a Jensen Shannon divergenceamong the current position of horse 'X_t' and best position of the horse 'X_{best}'.

3.2.3.2 Exploitation Phase

During the exploitation phase, every horse in herd changes its position towards best known position. In this stage, randomness movement is avoided to enable a more focused and accurate local search.

$$X_{t+1} = X_{best} + R_2 * 0.5|X_{best} - X_t| \quad (24)$$

Where, X_{t+1} indicates updated position of the horse, X_{best}represents current best location of the horse, R_2 denotes a random number (0,1), $0.5|X_{best} - X_t|$ represents a Jensen Shannon divergence, the current position 'X_t and 'X_{best} represents a best position of the horse. After updating the horse position during the exploration or exploitation, the algorithm again verifies the fitness.

$$FV = \begin{cases} f(X_{t+1}) > f(X_{best}) \ ; & X_{t+1} \ is\ optimal \\ Otherwise \ ; & X_{best}\ is\ optimal \end{cases} \quad (25)$$

Where, FVrepresentsa fitness verification results, If the newly updated position 'X_{t+1}' has a better fitness value than the best 'X_{best}', it returned newly updated position X_{t+1} as an optimal. Otherwise the best 'X_{best}' is chosen as optimal. The process gets iterated until reached the predetermined maximum iterations. As a result, the horses optimal position representing the best weight to reduce error rate in the crop prediction.

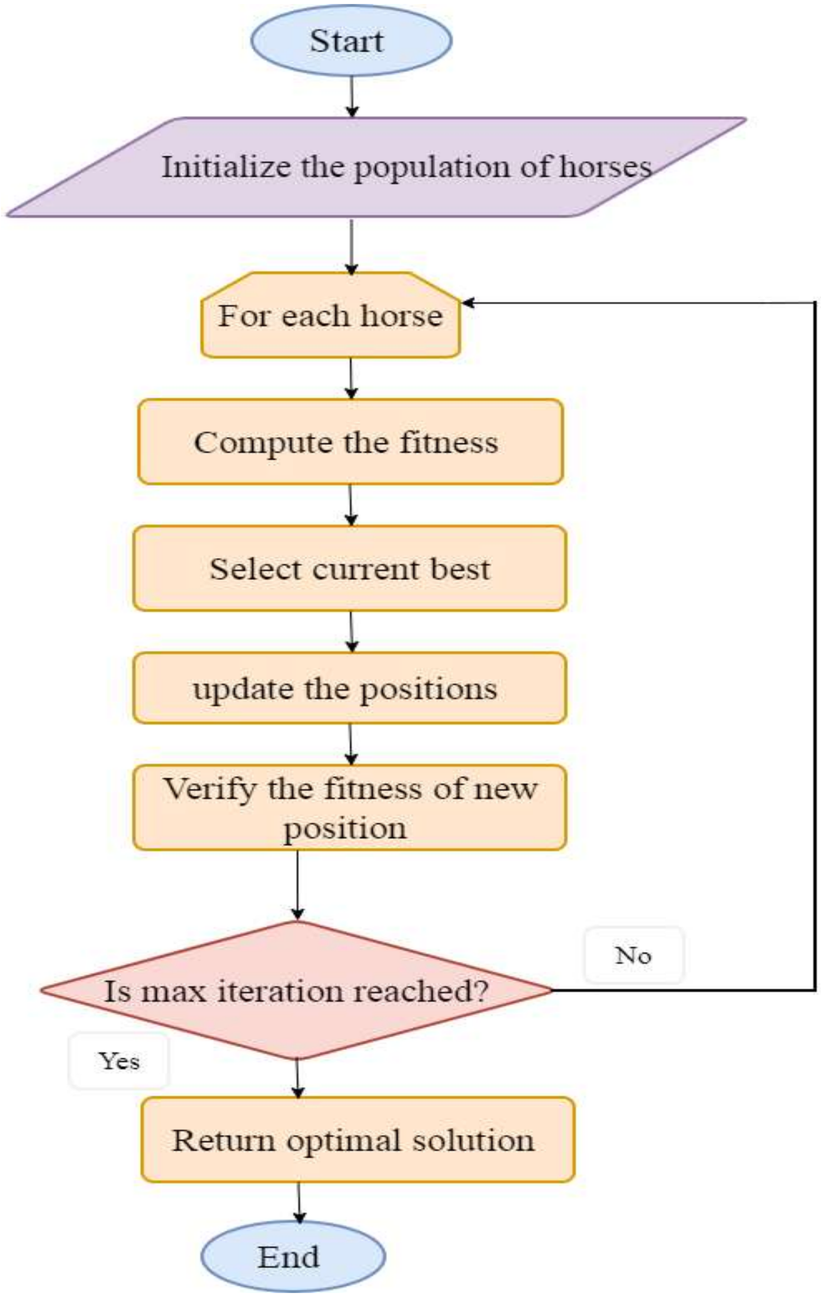


Figure 3. Flow Chart Of Horse Herd Optimization

Figure 3given above illustrates the flow chart of the horse herd optimization algorithm. Initially, a population of horses (weights) is generated within a search space. Each horse in this population is then evaluated using fitness function. Current best horse (weight) is chosen among the others for selecting the optimal along with the two behavioral approaches of the horse herd algorithm such as exploration and exploitation. Also, the current positions of horse are updated.With updating the horse position, the fitness is verified for new position of horse. This process repeated until the maximum iterations get reached. Finally, the optimal weight of deep learning network model is determined that minimizes the error function and allows the model to achieve the best possible performance on the classification task. Finally, the accurate crop classification results are obtained at output layer with softmax activation function.

$Y = \sigma_{soft}(h_t)$ (26)

Where,Yindicates a final classification output, $h_t$representsa output of hidden layer, softmax activation function ' σ_{soft}' at the output layer for providing the multi classoutput. In this way, Convolutional Deep Belief Network accurately performs the crop prediction enabling reliable recommendations for future agricultural planning and decision-making. The algorithmic process of classifier model is given below.

//**Algorithm 1: P**olynomial **Re**gressive**M**omentum **O**ptimized **D**iscriminative AI

Input: **Dataset '$\boldsymbol{DS}$', features**$A_1, A_2, \dots, A_m$ and data samples $TS_1, TS_2, \dots, TS_n$
Output: Improve the disease prediction accuracy

Begin
// Data acquisition
Step 1: Collect number of data samples $TS_1, TS_2, \dots, TS_n$ **a**nd features $A_1, A_2, \dots, A_m$ from the dataset
// Data processing
Step 2: Input features $A_1, A_2, \dots, A_m$ and data samples $TS_1, TS_2, \dots, TS_n$given to visible layer
Step 3: Transform input to RBM layer 1
Step 4: For each data samples TS_i**do**
Step 5: Compute the neuron activation probability using (1)
Step 6:End For
// Data Pre-processing ------Hidden layer 1
Step 7: For each data samples TS_i**do**
Step 8: Perform missing data imputation using (3)
Step 9: End for
Step 10: For each data samples TS_i
Step 11: Compute the Extreme normalized residual test '$ENRT$' using (5)
Step 12: If $(arg\ max\ ENRT)$ **then**
Step 13: Detect outlier data samples
Step 14: else
Step 15: Detect normal data samples
Step 16: End if
Step 17: End for
// feature selection ------ RBM layer 2
Step 18: For each preprocessed data samples do
Step 19: Perform feature selection
Step 20: Measure Roger similarity index using (9)
Step 21: $\boldsymbol{if}$ $(SI > HT)$ **then**
Step 22: Features said to be relevant and selected using (11)
Step 23: elseif $(SI < HT)$ **then**
Step 24: Features said to be irrelevant and removed using (10)
Step 25: End if
Step 26: End for
// classification------ RBM layer 3
Step 27: For each selectedfeature with training data samples
Step 28: For each testing data samples
Step 29: Measure the Chatterjee correlation coefficient 'CRC'using (12)
Step 30: $\boldsymbol{if}$ $(CRC = 1)$**then**
Step 31: Data samples classified into particular class
Step 32: **End if**
Step 33: End for
Step 34: End for
// Fine tuning
Step 35: For each classification result
Step 36: Compute the error 'CFEusing (17)
Step 37: Update the weight using (18)
Step 38: Initialize the population of weights using (21)
Step 39: For eachW_k
Step 40: Compute the fitnessusing (22)
Step 41: Identify current best weight with better fitness
Step 42: While ($\boldsymbol{t} < t_{max}$) do
Step 43:for each weight in population **do**
Step 44: if exploration phase then
Step 45: Update the positions 'X_{t+1}' using equation (23)
Step 46: Else if exploitation phase then
Step 47: Update the positions X_{t+1}' using equation (28)
Step 48: End if
Step 49: End for

Step 50: Verify fitness for updatedpositions using (25) **Step 51:** **if** $\left(f(X_{t+1}) > f(X_{best})\right)$**then** **Step 52:** X_{t+1} considered as best optimal solution **Step 53:** **else** **Step 54:** X_{best}considered as optimal solution **Step 55:** **End if** **Step 56:** Increment **t= t+1** **Step 57:** **Go to step 42** **Step 58:** **End while** **Step 59**: Return optimal weight **Step 60:** Obtain final classification results using (26)**------(output layer)** **End**

Algorithm 1 given above describes a different process of crop prediction for efficient recommendation with minimal error rates using PREMOD-AI model. The number of input climate and soil data are collected from the dataset. Then the input data samples are given to the visible layer of proposed deep learning network architecture. For each input data sample, neuron activation probability in measured along with the weights and biases to the RBM layer. The data then progresses through successive hidden layers, where data preprocessing is carried out for handling missing data and outlier data removal within the input dataset. In the next RBM layer, significant features are selected by employing the hockey-stick regression from the dataset to decrease the dimensionality of the dataset. Finally, Classification is carried out in the successive layer with the selected relevant features using the correlation coefficient, which measures the similarity between the training and testing samples to enhance crop prediction accuracy. After classification, a fine-tuning phase is employed using the horse herd optimization algorithm. Initially, a population of horses (representing weights) is spread across the search space. The fitness of these horses is evaluated based on the classification error rate, and their positions are updated iteratively. This process is repeated until a reaches the maximum number of iterations. After the algorithm convergence, the horse herd optimization algorithm identifies optimal weight values that reduce the classification error. Finally, the accurse multi class crop prediction results are observed at the output layer by applying the Softmax function, which further increases the classification accuracy.

4. EXPERIMENTAL SCENARIO

In this subdivision, experimental assessment of proposed model PREMOD-AI model and existingIDBN [1], GTODL-CRYPM [2], and TabNet [4] areimplemented in Python high level programming language. In order to conduct the experiment, crop recommendation dataset is collected from https://www.kaggle.com/datasets/irakozekelly/crop-recommendation-dataset. For the experimental analysis, the crop prediction is employed for precise crop recommendation. It combines soil characteristics, climatic conditions, and region-specific agricultural data. The dataset contains testing and training folders with consist of 18,079 instances with eight attributes. These attributes include a unique record identifier, key soil nutrients such as nitrogen (N), phosphorus (P), potassium (K), and pH level, along with environmental variables including temperature and rainfall, and the corresponding crop type. Detailed explanations of these features are presented in Table 2and Table 3.For experimental purposes, the number of data samples is varied incrementally, ranging from 1800 to 18000.During the hardware and software configuration, the entire experiment is conducted in an Intel Core i5- 6200U CPU @ 2.30GHz 4 cores with 8 Gigabytes of RAM.Cross-validation is used to compute the performance of the model ability to carry out unseen data. By using cross-validation, the dataset is divided into two sets such as training and testing. Most data samples (70%) were used for training and the remaining (20% and 10%) were taken for testing and validation. In our work, the 10- fold cross-validation is utilized for measuring the results. Hyperparameter settings are described in table 4.

Table 4 Hyperparameter description

Hyperparameter	**Description**
Number of layers	One visible, three hidden (RBM 3) and output layer
RBM layer 1	Data preprocessing performed using polynomial proximal regressive imputation and Extreme Normalized Residual Test Statistic
RBM layer 2	Roger similarity index used
RBM layer 3	Chatterjee correlation coefficient employed
Learning rate	0.001
Number of epochs	100
Batch size	64
Optimizer	Adam

Activation function	Softmax
Training: Testing: Validation	70:20:10

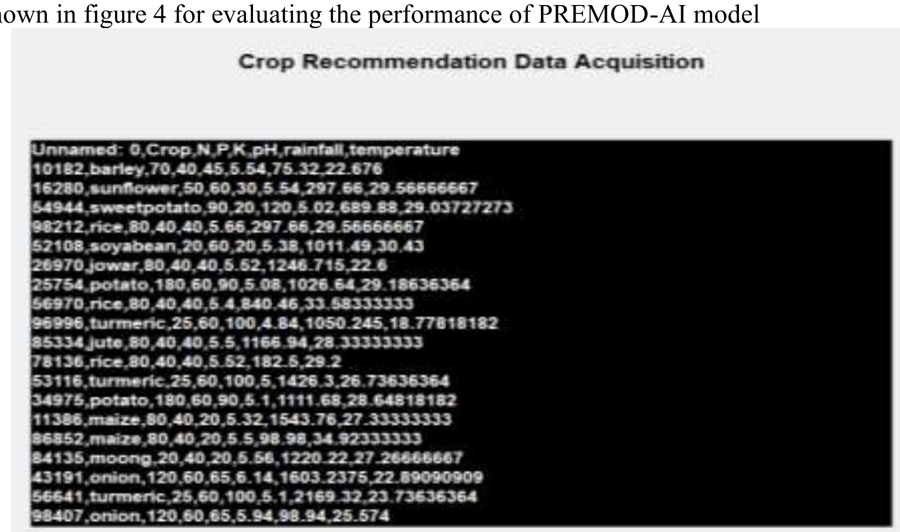
4.1 Implementation Details

In this study, we developed PREMOD-AI for crop recommendation.

- The PREMOD-AI method comprises of an input layer, three hidden layers and an output layer.
- We compared our PREMOD-AI method to model to existing IDBN [1], GTODL-CRYPM [2] and TabNet [4] using datasets to validating the results.
- Initially, distinct 18000 sample X-ray images were obtained from crop recommendation dataset. The crop recommendation Dataset containing data samples were initially positioned in the input layer.
- Second in the first hidden layer, data preprocessing is performed by using polynomial proximal regressive imputation and Extreme normalized residual test for handling the missing data and removing the outliers.
- Third, in the second hidden layer, feature selection is carried out employing Hockey-stick regression via Roger similarity statistical index for selecting the relevant features and eradicating the immaterial features.
- Fourth, in the third hidden layer, Chatterjee correlation coefficients are used for precise crop prediction. Stochastic adaptive moment algorithm and horse herd optimization are applied for hyperparameters by minimizing the error rate.
- Finally, in the output layer, using Softmax as the activation function, accurse multi class crop prediction results are obtained for prediction of crop recommendation.

4.2 Implementation Results

The PREMOD-AI model is comprehensively observed to estimate its effectiveness in crop prediction. The evaluation includes several core stages, including data collection, preprocessing, feature selection, and classification, utilizing the crop recommendation dataset taken from Kaggle repository. First, the 18000 data sample data are collected from the dataset is shown in figure 4 for evaluating the performance of PREMOD-AI model


Crop Recommendation Data Acquisition

Unnamed: 0,Crop,N,P,K,pH,rainfall,temperature
10182,barley,70,40,45,5.54,75.32,22.676
16280,sunflower,50,60,30,5.54,297.66,29.56666667
54944,sweetpotato,90,20,120,5.02,689.88,29.03727273
98212,rice,80,40,40,5.66,297.66,29.56666667
52108,soyabean,20,60,20,5.38,1011.49,30.43
26970,jowar,80,40,40,5.52,1246.715,22.6
25754,potato,180,60,90,5.08,1026.64,29.18636364
56970,rice,80,40,40,5.4,840.46,33.58333333
96996,turmeric,25,60,100,4.84,1050.245,18.77818182
85334,jute,80,40,40,5.5,1166.94,28.33333333
78136,rice,80,40,40,5.52,182.5,29.2
53116,turmeric,25,60,100,5,1426.3,26.73636364
34975,potato,180,60,90,5.1,1111.68,28.64818182
11386,maize,80,40,20,5.32,1543.76,27.33333333
86852,maize,80,40,20,5.5,98.98,34.92333333
84135,moong,20,40,20,5.56,1220.22,27.26666667
43191,onion,120,60,65,6.14,1603.2375,22.89090909
56641,turmeric,25,60,100,5.1,2169.32,23.73636364
98407,onion,120,60,65,5.94,98.94,25.574

Figure 4 sample data collection from the dataset

The class distribution of 18,079 instances labeled data samples, which are distributed among 40 distinct categories. This class distribution ensures that the model receives a representation of data samples.

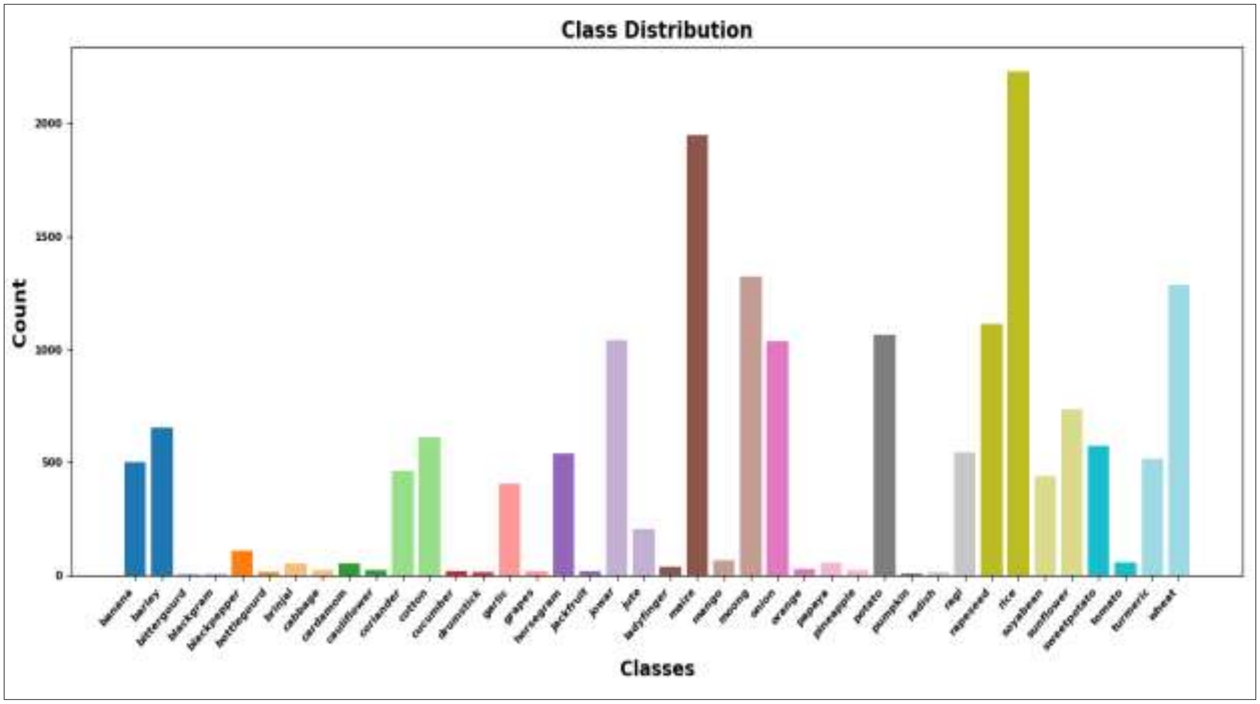


Figure 5. Class Distribution Of Dataset

The bar chart displays the class distribution of 40 categories of crop. On the x-axis, the 40 class labels are represented. The y-axis indicates the count of samples for each class.

After data collection, preprocessing steps is employed, including handling missing values and eliminating outliers. Initially, the total dataset size is 782 KB, which was reduced to 776 KB after preprocessing. The preprocessing results are shown in figure 6.

```
Unnamed: 0,Crop,N,P,K,pH,rainfall,temperature
10182,barley,70,40,45,5.54,75.32,22.676
16280,sunflower,50,60,30,5.54,297.66,29.56666667
54944,sweetpotato,90,20,120,5.02,689.88,29.03727273
98212,rice,80,40,40,5.66,297.66,29.56666667
52108,soyabean,20,60,20,5.38,1011.49,30.43
26970,jowar,80,40,40,5.52,1246.715,22.6
25754,potato,180,60,90,5.08,1026.64,29.18636364
56970,rice,80,40,40,5.4,840.46,33.58333333
96996,turmeric,25,60,100,4.84,1050.245,18.77818182
85334,jute,80,40,40,5.5,1166.94,28.33333333
78136,rice,80,40,40,5.52,182.5,29.2
53116,turmeric,25,60,100,5.0,1426.3,26.73636364
34975,potato,180,60,90,5.1,1111.68,28.64818182
11386,maize,80,40,20,5.32,1543.76,27.33333333
86852,maize,80,40,20,5.5,98.98,34.92333333
84135,moong,20,40,20,5.56,1220.22,27.26666667
43191,onion,120,60,65,6.14,1603.2375,22.89090909
56641,turmeric,25,60,100,5.1,2169.32,23.73636364
98407,onion,120,60,65,5.94,98.94,25.574
89586,moong,20,40,20,5.6,207.55,26.43333333
43620,cabbage,100,65,70,6.62,1111.68,28.64818182
```

Figure 6. Data Preprocessing

The PREMOD-AI model focuses on identifying relevant input features that enhance the accuracy of crop recommendation. To achieve this, itHockey-stick indexed regression to select 6 key features which has similarity score close to ‘1’ from the 8 features. The heatmap and selected features demonstrated in figure 7 and 8.

Figure 7Heatmap Similarities Between The Features

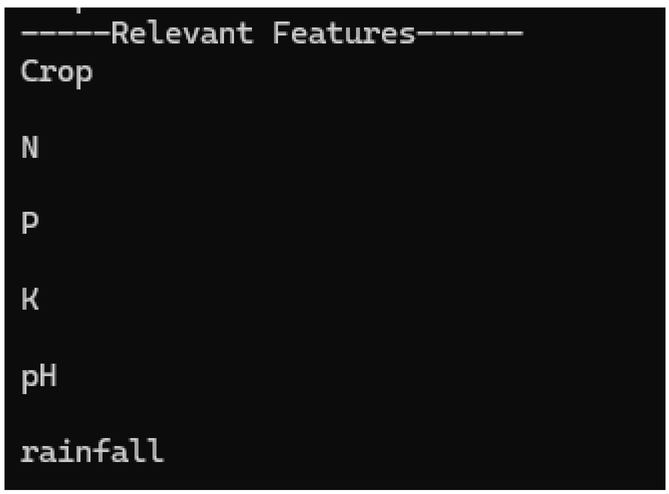


Figure 8 Selected 6 features

Finally, the crop classification is carried out in figure 9 using Chatterjee correlation coefficients. The method is trained using 6 selected relevant features from data samples to classify the different crops, thereby minimizing the classification error.

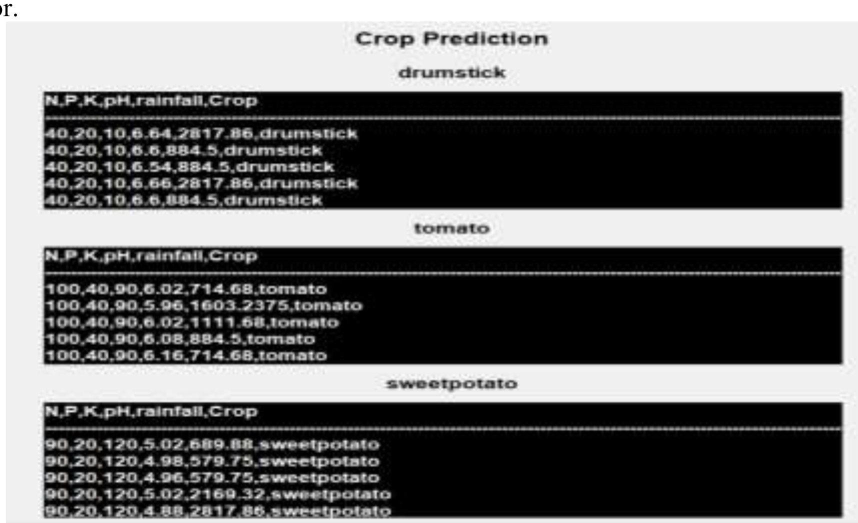


Figure 9 Crop Prediction

5. PERFORMANCE COMPARISON ANALYSIS

In this section, the performance of the proposed PREMOD-AI model is evaluated and compared with existing approaches IDBN [1], GTODL-CRYPM [2] andTabNet [4]. The assessment is carried out using several metrics, including accuracy, precision, recall, F-score, RMSE and crop recommendation time. The results are summarized in tables and illustrated through graphs to facilitate comparison.

5.1 Evaluation metrics

The evaluation measures employed to evaluate multiclass classification in crop prediction are as follows.

Accuracy: It refers to the proportion of data samples that are correctly identified as specific crop to the number of samples taken as input from the dataset. The metric is computed using the following formula

$$Accuracy = \left(\frac{TP+TN}{TP+\boldsymbol{TN}+FP+FN}\right) * 100 \qquad (27)$$

Where, True Positives (TP) refers to data samples correctly detected specific crop type. while True Negatives (TN) represents refer to data samples where the model correctly recognizes that a particular crop type is not present. False Positives (FP) occur when samples is incorrectly detected predicts a crop type resulting in false alarms. Finally, False Negatives (FN**)** represents the data samples where the model correctly recognizes that a particular crop type is not present. The accuracy is measured in percentage (%).

Precision: It is measured by the proportion of correctly classified positive data samples (true positives) out of data samples predicted as positive, which includes both true positives and false positives. The precision of crop classification is expressed as follows,

$$\boldsymbol{Precision = \left(\frac{TP}{TP+FP}\right) * 100} \qquad (28)$$

Where, TPrepresents the true positive, FPindicates the false positive.

Recall: it also referred to as sensitivity, computes the ability of a model to correctly identify specific crop types. It is mathematically calculated by measuring the number of true positive predictions by the sum of true positives and false negatives. The formula for recall is expressed as follows:

$$\boldsymbol{Recall = \left(\frac{TP}{TP+FN}\right) * 100} \qquad (29)$$

Where,TPrepresents the true positive, FNindicates the false negative.

F1score: It alsoknown F-measure measured as the harmonic mean of both precision as well as recall. It is mathematically computed as follows

$$F1 - Score = 2 * \left(\frac{Precision*Recall}{Precision+Recall}\right) \qquad (30)$$

Root Mean Square Error: It measured as the ratio of the difference between the actual predicted outcomes and observed results to the total number of data samples. Therefore, root mean square error is calculated by using following expression,

$$RMSE = \left[\sqrt{\frac{\left(Y_{actual}-Y_{prediction}\right)^2}{n}}\right] \qquad (31)$$

Where, $RMSE$ represents a root mean square error, Y_{actual}symbolizes the actual prediction results$Y_{prediction}$ represents a predicted output.

Matthews Correlation Coefficient (MCC): It is a balanced evaluation metric that considers true positives, true negatives, false positives, and false negatives in crop prediction. It provides a particular score to price the quality of multiclass crop prediction.

$$\boldsymbol{MCC} = \frac{(TP*TN)-(FP*FN)}{\sqrt{(TP+FP)(TP+FN)\,(TN+FP)(TN+FN)}} \qquad (32)$$

Where, MCC denotes a Matthews Correlation Coefficient, true positive 'TP' and the false negative 'FN', and false positive rate 'FP', true negative 'TN'.

R² score: It also known as the coefficient of determination is a numerical metric used to calculate the performance of crop prediction outcomes. It is also defined as the ratio of the residual sum of squares between the actual and predicted crop to the total sum of squares.

$$R^2 = 1 - \frac{\left(Y_{Actual}-Y_{prediction}\right)^2}{(\hat{Y}-Y_{Actual})^2} \qquad (33)$$

Where, $\boldsymbol{R^2}$denotes coefficient of determination, $\boldsymbol{Y_{Actual}}$ denotes the actual results of crop prediction,$\boldsymbol{Y_{prediction}}$ denotes a predicted results of crop, $\boldsymbol{\widehat{Y}}$ denotes a mean of actual value.

Crop recommendation time: This metric represents the total amount of time required by the algorithm to find the most suitable crop recommendations across all data samples. It is quantitatively measured using the following mathematical expression**,**

$$CPRT = \sum_{i=1}^{n} AS_i * T\,[SCC] (34)$$

Where, $CPRT$represents a crop recommendation time, AS_i indicates a number of instances or data samples, $T\ [SCC]$indicates a time for crop prediction. It is measured in milliseconds (ms).

5.2 EVALUATION OF RESULTS

This section outlines the experimental assessments obtained using the proposed PREMOD-AI model against other deep learning algorithms. These results obtained explicitly illustrate that the PREMOD-AI model outperforms the other algorithms evaluated.

Table 4 comparison of accuracy versus data samples

Number of data samples	Accuracy (%)			
	Proposed PREMOD-AI	Existing IDBN [1]	Existing GTODL-CRYPM [2]	Existing TabNet [4]
1800	95.55	92.22	90	88.88
3600	95.53	92.2	89.56	88.43
5400	95.45	92.18	89.42	88.12
7200	94.88	92.15	89.41	88.09
9000	94.74	92.32	88.88	87.76
10800	95.68	92.63	89.45	88.09
12600	95.71	92.45	89.23	88.15
14400	95.66	92.33	89.41	88.37
16200	95.37	92.41	89.63	88.65
18000	95.35	92.11	89.42	88.31

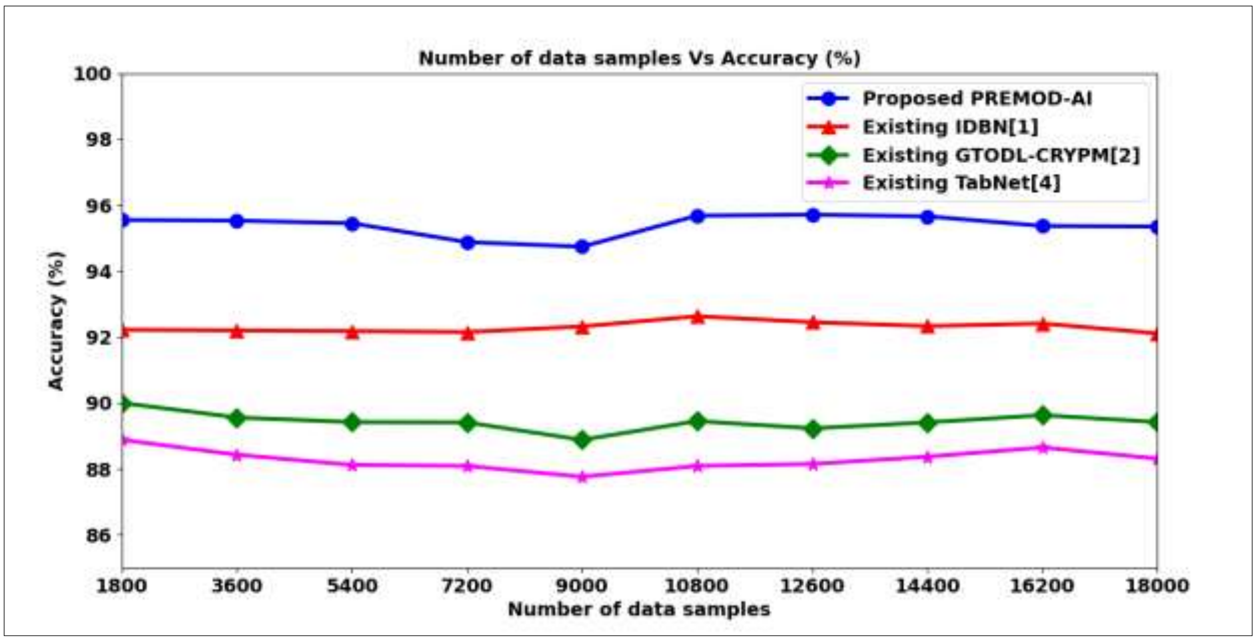


Figure 10 Graphical illustration of accuracy

Table 4 and Figure 10 present a graphical depiction of crop prediction accuracy results observed by applyingfour methods namely PREMOD-AI model is evaluated and compared with existing approaches IDBN [1], GTODL-CRYPM [2] and TabNet [4]. As revealed in Figure 10, the horizontal axis represents the number of data samples, collected from the dataset ranging from 1800 to 18000, while the vertical axis represents the crop prediction accuracy. Among the four methods, the outcomes of the PREMOD-AI model demonstrateconsiderable improvement compared to the existing deep learning methods. For example, when considering1800 data samples in the first run, the PREMOD-AI model achieved an accuracy of 95.55%. In contrast, the prediction accuracy of the existing models [1] [2] and [4] were observed to be92.22% and 90% and 88.88% respectively. Similarly, tenvarious sets of results were evaluated for each deep learning method along with the different number of data samples. The overall observed results of the PIPR-TOTL model were then compared to those of the existing methods. The comparative analysis of the ten results demonstrated that the crop prediction accuracy of the PREMOD-AI model improved by 4%, 7% and 8% over

[1] [2] and [4], respectively. This improvement is achieve due to the application of Chatterjee correlation coefficient which analyzes the training and testing data samples in the hidden layer of the deep belief model resulting it enhances the accuracy. Based on correlation coefficient, accurately are predicted with better accuracy. Moreover, the fine-tuning process of the proposed deep learning model utilizes horse herd optimization to reduce classification error by increasing the true positive and true negative rates, while decreasing the false positives and false negatives in the more accurate crop predictions.

Table 5 comparison of Precisionversus data samples

Number of data samples	Precision (%)			
	Proposed PREMOD-AI	Existing IDBN [1]	Existing GTODL-CRYPM [2]	Existing TabNet [4]
1800	95.61	92.92	91.22	90.39
3600	95.57	92.85	91.18	90.08
5400	95.45	92.74	91.12	89.78
7200	95.36	92.36	90.26	89.37
9000	95.88	92.45	90.57	89.44
10800	95.36	92.37	90.63	89.65
12600	95.41	92.21	90.44	89.23
14400	95.05	92.45	90.33	89.08
16200	95.74	92.63	90.08	89
18000	95.33	92.11	90.64	89.23

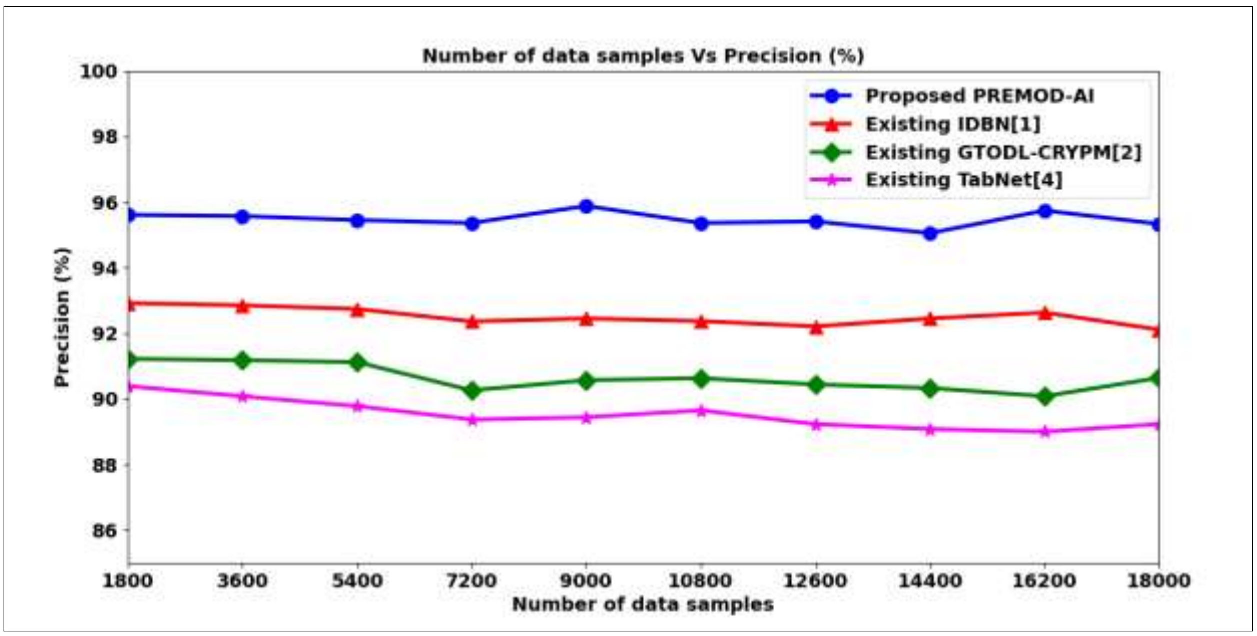


Figure 11 Graphical illustration of precision

Table 5 and Figure 11 illustrate the graphical illustration of precision regarding the number of data samples collected from the dataset ranges from 1800 to 18000. The overall performance of precision in crop recommendation is calculated using four deep learning methods namely PREMOD-AI model, IDBN [1], GTODL-CRYPM [2] and TabNet [4].The horizontal axis represents the number data samples, while the perpendicular axis of the graph designates the overall precision performance. Let us consider 18000 samples for computing the precision. In order to apply the PREMOD-AI model achieved a precision of 95.61%, while the existing methods [1] [2] and [4] observed to be 92.92%, 91.22% and 90.39%, respectively. Similarly, different precision results were observed regarding various counts of input data samples. Finally, the observed results of PREMOD-AI model are compared to the existing methods [1] and [2]. The average of ten comparison results indicates that the PREMOD-AI model shows an improvement in precision of approximately 3% over method [1] and around 5% over method [2] and 7% over method [4].This enhanced performance of PREMOD-AI model is achieved by integrating the convolutional deep belief network in combination with the Chatterjee correlation coefficient for data analysis based on feature assessment.

Moreover, Fine-tuning layers of the deep learning model is donethrough the integration ofHorse herdoptimization algorithm to minimize the training and validation errors. This approach efficiently enhances performance of precision by increasing the true positive rate while reducing false positives in crop prediction.

Table 6 comparison of recallversus data samples

Number of data samples	Recall (%)			
	Proposed PREMOD-AI	Existing IDBN [1]	Existing GTODL-CRYPM [2]	Existing TabNet [4]
1800	97.32	94.59	92.85	92
3600	97.25	94.46	92.78	91.68
5400	97.11	94.33	92.36	91.53
7200	97.08	94.23	92.44	91.58
9000	97.36	94.35	92.36	91.42
10800	97.25	94.28	92.55	91.66
12600	97.09	94.48	92.69	91.75
14400	97.33	94.66	92.78	91.82
16200	97.28	94.35	92.62	91.72
18000	97.22	94.41	92.77	91.65

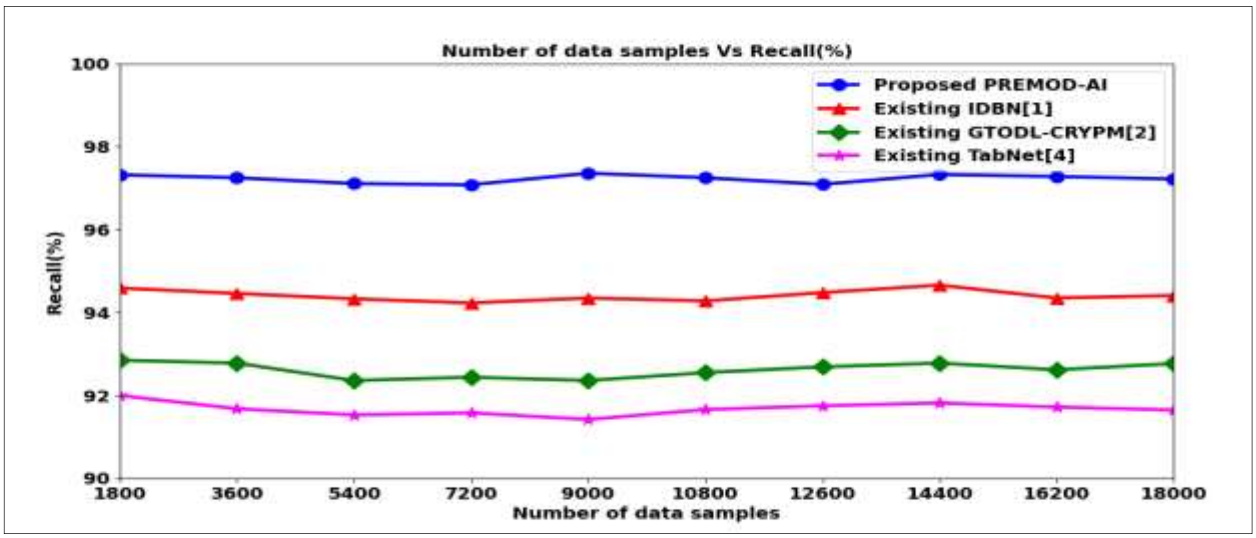


Figure 12 Graphical illustration of recall

Table 6 and Figure 12 illustrate an overall comparison analysis of recall performance along with the number of climate and soil related data samples collected from the dataset. The PREMOD-AI model exhibits higher recall performance in crop prediction compared to the existing approaches [1] [2] and [4]. For example, with 1800 data samples, PREMOD-AI model achieved a recall of 97.32%, while methods [1] [2] and [4] observed to be 94.59%, 92.85% and 92%, respectively. Similarly various performance results were observed according to different counts of input samples. The experimental analysis shows that the average recall of PREMOD-AI model improved by approximately 3% compared to [1] and 5% compared to [2] and 6% compared to [4].The superior performance of the PREMOD-AI model is achieved due to the fine-tuning process applied within the deep learning framework. By utilizing a convolutional deep belief approach, the model minimizes the classification error between predicted and actual outcomes in multiclass output through optimal weight adjustment using the horse herd optimization algorithm. This procedure iterated continue until the error gets minimized, resulting in lesser false negatives and an increase in true positives for precise crop prediction in agriculture sector. .

Table 7Comparison of F1scoreversus data samples

Number of data samples	F1 score (%)			
	Proposed PREMOD-AI	Existing IDBN [1]	Existing GTODL-CRYPM [2]	Existing TabNet [4]
1800	96.45	93.74	92.02	91.18

3600	96.40	93.64	91.97	90.87
5400	96.27	93.52	91.73	90.64
7200	96.21	93.28	91.33	90.46
9000	96.61	93.39	91.45	90.41
10800	96.29	93.31	91.57	90.64
12600	96.24	93.33	91.55	90.47
14400	96.17	93.54	91.53	90.42
16200	96.50	93.48	91.33	90.33
18000	96.26	93.24	91.69	90.42

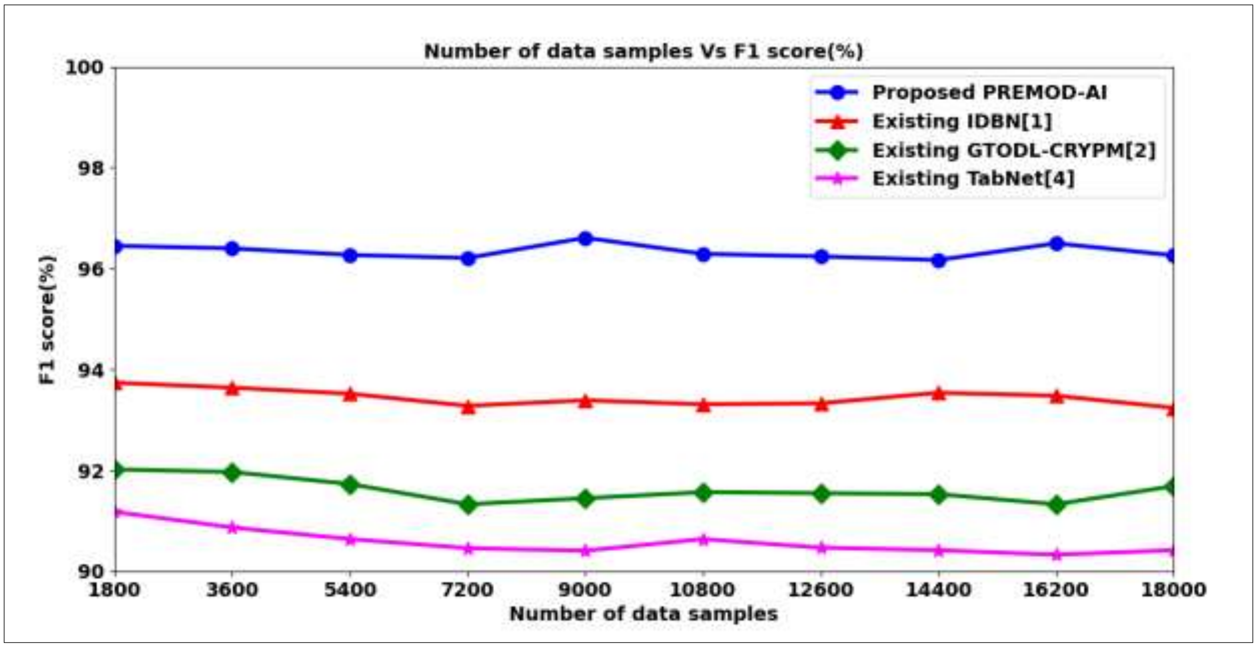


Figure 13 Graphical illustrations of F1 score

Table 7 and Figure 13 present a performance assessment of F1 scores obtained using four methods namely the proposed PREMOD-AI model, IDBN [1], GTODL-CRYPM [2] and TabNet [4]. To facilitate a widespread analysis, varying counts of data were used as input, ranging from 18,000 to 18,000. For each input data, the F1 score was calculated for all four models, and the results are summarized. The results designate that the proposed PREMOD-AI model consistently outperforms conventional deep learning models in terms of F1 score. Each method was estimated over ten independent runs to achieve better F1 measurements. The analysis demonstrates that PREMOD-AI modelachieves superior F1 score performance across all experiments. On average, over the ten runs, PREMOD-AI model showed an improvement of approximately 3% over method [1] and 5% over method [2] and 6% over method [4]. The improved performance is achieved due to the efficient application of a deep learning framework, which enhances the F1-score by achieving an optimal balance between precision as well as recall. This balanced performance enables more reliable and consistent classification of multiple crops in agriculture sector.

Table 8Comparison of RMSEversus Data samples

Number of data samples	**RMSE**			
	Proposed PREMOD-AI	**Existing IDBN [1]**	**Existing GTODL-CRYPM [2]**	**Existing TabNet [4]**
1800	0.104	0.183	0.235	0.262
3600	0.074	0.13	0.174	0.192
5400	0.061	0.106	0.143	0.161
7200	0.060	0.092	0.124	0.14
9000	0.055	0.080	0.117	0.129
10800	0.041	0.070	0.101	0.114
12600	0.038	0.067	0.095	0.105

14400	0.036	0.063	0.088	0.096
16200	0.036	0.059	0.081	0.089
18000	0.034	0.058	0.078	0.087

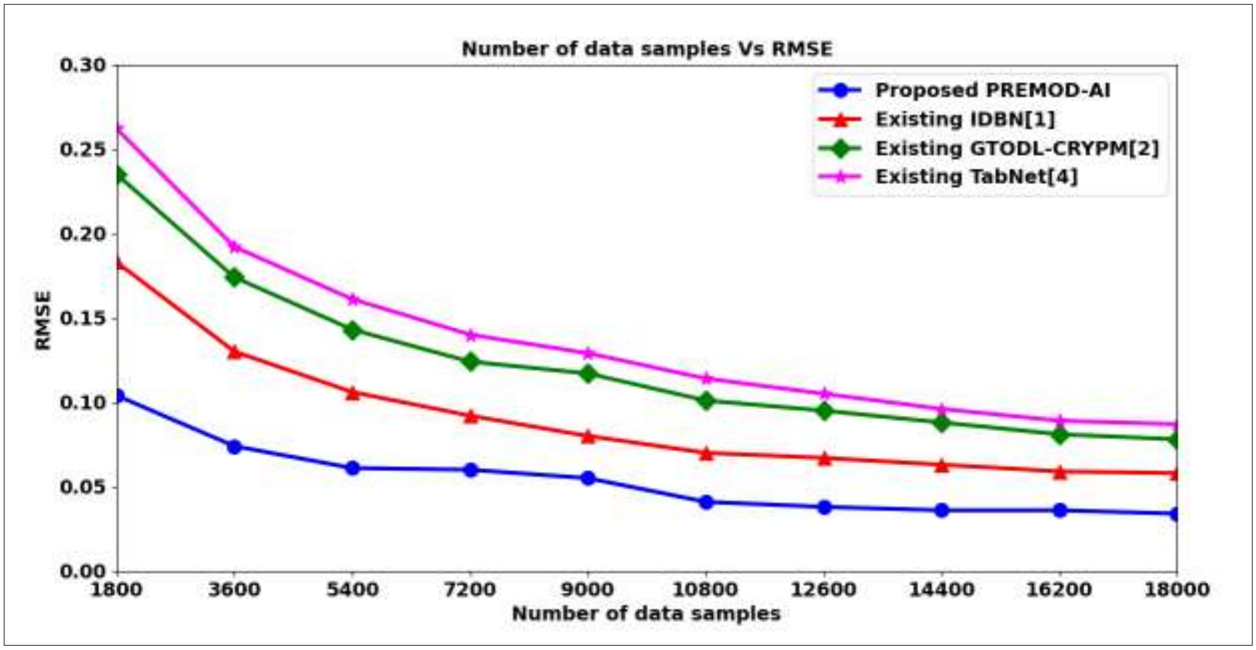


Figure 14 Graphical illustrations of RMSE

Table 8 and Figure 14 illustrate theperformanceoutcomes of RMSE in multiple crop classification versus the number of data samples, ranging from 1800 to 18000. Four methods were employed for computing the RMSE in crop recommendation namely the PREMOD-AI model and existing deep learning model IDBN [1], GTODL-CRYPM [2] and TabNet [4].As exposed in Figure 14, the horizontal axis represents the number of data samples, while the vertical axis indicates the RMSE. The experimental results indicate that the proposed PREMOD-AI model consistently achieves minimal RMSE values compared to the three existing approaches for crop classification. Let us consider an initial experiment with 1800 data samples, the PREMOD-AI model recorded an RMSE of 0.104, whereas methods [1] [2] and [4] produced RMSE values of 0.183, 0.235 and 0.262, respectively. Overall, the results reveal that the PREMOD-AI model reduces overall performance of RMSE by 40% in comparison with [1] and by 57% relative to [2] and 61% in comparison with [4]. This performance improvement is mainly achieved due to the model's effective use of the stochastic adaptive moment algorithm for fine-tuning the hyperparameters of the convolutional deep belief network classifier. By reducing the error at the output layer, horse herd optimization strategy considerably improves the accuracy of crop prediction.

Table 9. Comparison of MCCversus Data samples

Number of data samples	MCC			
	Proposed PREMOD-AI	Existing IDBN [1]	Existing GTODL-CRYPM [2]	Existing TabNet [4]
1800	0.905	0.834	0.786	0.761
3600	0.904	0.832	0.785	0.758
5400	0.903	0.828	0.784	0.756
7200	0.902	0.833	0.782	0.753
9000	0.905	0.835	0.786	0.755
10800	0.906	0.837	0.788	0.759
12600	0.907	0.834	0.789	0.763
14400	0.906	0.833	0.786	0.761
16200	0.905	0.837	0.785	0.752
18000	0.904	0.838	0.784	0.75

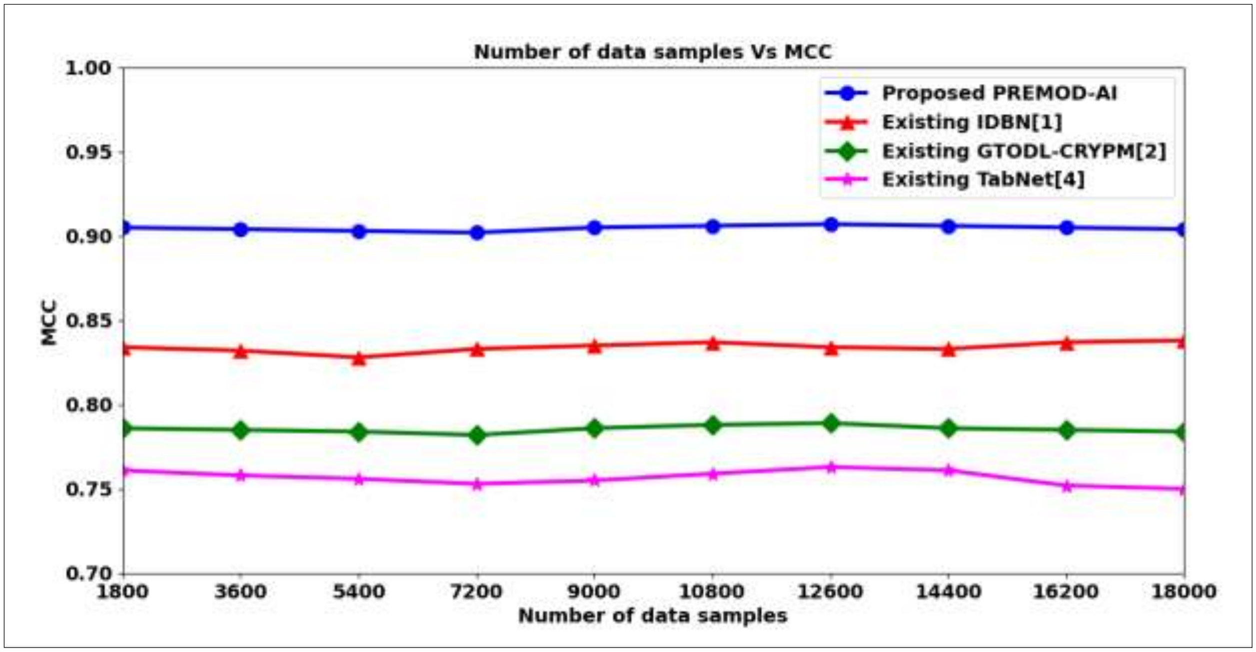


Figure 15 Graphical illustrations of MCC

Table 9 and Figure 15 illustrate the Matthews Correlation Coefficient (MCC) values obtained using the proposed PREMOD-AI model in comparison with existing classification methods. The MCC is a comprehensive metric that accounts for true and false positives and negatives, making it specifically useful for estimating the classification performance. In the figure, the x-axis represents the number of data samples, while the y-axis shows the corresponding MCC values calculated using 18,000 samples. The results indicate that the PREMOD-AI model consistently outperforms the other approaches, achieving higher MCC scores across all sample sizes. Specifically, consider the 1800 samples, thePREMOD-AI model achieved an MCC of 0.905, whereas method [1] achieved 0.834 and method [2] scored 0.786 and method [4] scored 0.761. These findings demonstrate the PREMOD-AI model ability to offer accurate crop classification results and efficiently reduce false predictions. The marked improvement in MCC underscores the robustness and reliability of the proposed PREMOD-AI model for crop prediction.The average of ten results indicate that the proposed PREMOD-AI model increases the MCC by approximately 8%, 15% and 20% compared to the existing techniques

Table 10. Comparison of R^2 scoreversus Data samples

Number of data samples	R^2 score			
	Proposed PREMOD-AI	Existing IDBN [1]	Existing GTODL-CRYPM [2]	Existing TabNet [4]
1800	0.926	0.896	0.885	0.874
3600	0.922	0.895	0.882	0.87
5400	0.921	0.887	0.881	0.868
7200	0.923	0.884	0.88	0.866
9000	0.925	0.883	0.874	0.863
10800	0.928	0.885	0.876	0.865
12600	0.926	0.882	0.872	0.861
14400	0.924	0.886	0.877	0.864
16200	0.926	0.884	0.875	0.862
18000	0.925	0.884	0.874	0.861

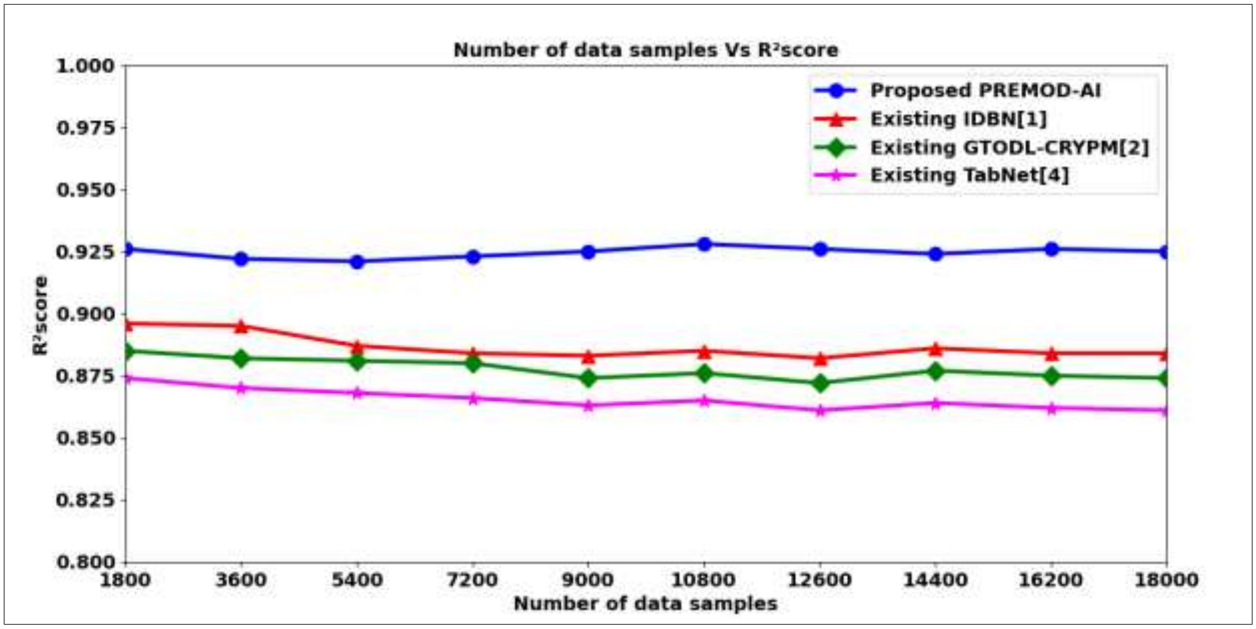


Figure 16 Graphical illustrations of R^2 score

Table 10 and Figure 16 show the performance of R^2 scores of the data sample size, which ranges from 1800 to 18,000. The x-axis represents the number of data samples, and the y-axis indicates the R^2 values obtained by each model. The proposed PREMOD-AI model is compared with three existing methods, [1] [2] and [4]. Across all sample sizes, the PREMOD-AI model consistently achieves higher R^2 scores. For example, at a sample size of 1800, the PREMOD-AI model attains an R^2 of 0.926, whereas methods [1] [2] and [4] reached 0.896, 0.885 and 0.874 respectively. Over ten independent trials, PREMOD-AI model shows an average improvement of approximately 4% over [1], 5% over [2] and 7% over [4]. This performance enhancement is due to the integration of the stochastic adaptive moment algorithm for fine-tuning the hyperparameters. By minimizing the prediction error at the output layer, horse herd optimization approach also significantly enhances the accuracy of crop prediction. In this context, an R^2 value closer to 1 indicates a strong predictive capability.

Table 11Comparison of crop recommendation Timeversus Data samples

Number of data samples	crop recommendation time (ms)			
	Proposed PREMOD-AI	Existing IDBN [1]	Existing GTODL-CRYPM [2]	Existing TabNet [4]
1800	54	64.8	70.2	80.5
3600	60.5	68.9	75.8	80.2
5400	65.8	75.8	80.5	83.4
7200	68.2	80.2	85.6	86.7
9000	75.8	83.6	88.6	90.3
10800	78.2	88.9	92.8	94.2
12600	85.6	92.6	97.5	103.5
14400	88.4	95.8	102.5	106.7
16200	90.2	98.2	109.6	112.6
18000	95.7	105.4	115.2	124.5

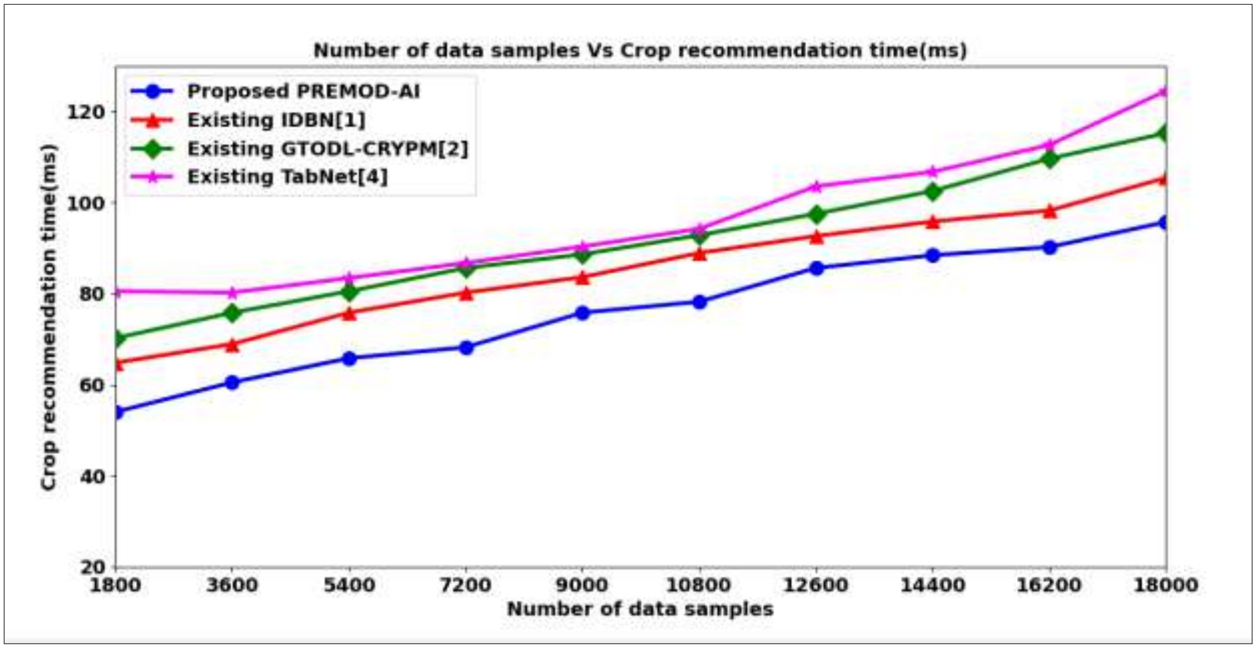


Figure 17 Graphical illustrations of crop recommendation time

Table 11 and Figure 17 illustrates the performance analysis of crop recommendation time using the proposed PREMOD-AI model and existing deep learning model IDBN [1], GTODL-CRYPM [2] and TabNet [4]. The detailed plots crop recommendation time against the number of input data samples, ranging from 1800 to 18000. As the number of data sample increases, the crop recommendation time for all four methods increases accordingly. The x-axis indicates the number of samples collected from the dataset, while the y-axis indicates the corresponding crop recommendation time. In an experiment with 18000 data samples, the PREMOD-AI model achieved a computation time of $54ms$, while methods [1] [2] and [4] observed a recommendation time of $64.8ms$, $70.8ms$and $80.5ms$respectively. Among four methods, the PREMOD-AI model constantly decreases the crop recommendation time by 11%, 17% and 21% compared to the three existing models. This effectiveness is achieved due to its pre-processing and feature selection process, which integrates into the layers of convolutional deep belief network architecture model. Polynomial proximal regressive imputation process is applied to handle missing data in a given dataset and Extreme normalized residual test applied for outlier removal. Moreover, the Hockey-stick indexed regression based feature selection process is carried out in PREMOD-AI model constantly to select the more relevant features to minimize to time consumption crop prediction. Overall, the analysis confirms that the PREMOD-AI model reduced computation time in crop recommendation.

5.3 Ablation study

The ablation study is carried out in Convolutional Deep Belief Networks(CDBN) to find out the contribution of preprocessing, feature selection and classificationimplemented to enhance overall performance. CDBN is carried out to perform efficient crop recommendation. CDBN is combined with the preprocessing (polynomial proximal regressive imputation and Extreme Normalized Residual Test Statistic),feature selection (Hockey-stick indexed regression), classification (Chatterjee correlation coefficients), and optimization (horse herd). The assessment is carried out to improve the crop recommendationaccuracy with minimum time consumption.

Table 12 Ablation Study

Methods/Name	**Crop recommendation dataset**							
	Accuracy (%)	**Precision (%)**	**Recall (%)**	**F-score**	**RMSE**	**Crop recommendation time (ms)**	**MCC**	**R² score**
CDBN	93.25	92.84	95.42	94.64	0.074	89.65	0.78	0.8
CDBN+ polynomial proximal regressive imputation,	93.78	92.96	95.83	94.78	0.071	86.68	0.82	0.83

Extreme Normalized Residual Test Statistic								
CDBN+ Hockey-stick indexed regression	94.05	93.42	96.15	94.89	0.068	83.45	0.85	0.84
CDBN+ Chatterjee correlation coefficient	94.32	93.65	96.33	95.09	0.064	80.32	0.89	0.86
CDBN+ horse herd optimization	94.65	93.87	96.54	95.21	0.057	77.85	0.86	0.88
Proposed model	**94.88**	**94.17**	**96.85**	**95.34**	**0.051**	**74.55**	**0.89**	**0.91**

Table 12 shows the ablation performance for crop recommendation dataset. CDBN model achieved accuracy of 93.25%respectively. The proposed PBOIDCL method attained higher accuracy of 94.88%, precision of 94.17%, and recall of 96.85% and F-score of 95.34% oncrop recommendation dataset. Also, the proposed PBOIDCL method of RMSE, time, MCC and R^2 score is reduced by 0.051, 74.55ms, 0.89 and 0.91. The experimental results show the efficiency of PREMOD-AI model than other methods.

5. DISCUSSION

Proposed PREMOD-AI is an artificial intelligence (AI) that helps farmers pick the best crops to grow. It evaluates soil, weather, and water data. This allows farmers to save money, increase crop yields, and use fewer resources. It provides fast, accurate crop recommendations for smart farming. It uses advanced algorithms to correctly identify the best crops to plant with higher accuracy.It gives advice in real-time, which helps farmers make quick choices. It lowers waste by using data to guide planting choices.The time complexity is very low by using AI.

From result analysis, the proposed PREMOD-AI attains strong crop recommendation classification performance with high accuracy, precision, recall and minimum time. The classification results are used for smart agriculture in real-time application. The proposed PREMOD-AI method helps the agriculture professionals for crop recommendation with minimum resources. PREMOD-AI method is used in crop monitoring systems for timely efficient crop recommendation. The key issue of PREMOD-AI method is not reduced the space complexity while collecting and storing data samples. PREMOD-AI method is not tested for remote monitoring systems.

6. CONCLUSION

In this article, a novel PREMOD-AI model is developed for accurate crop recommendation. It helps the farmer to find the appropriate crop and its expected yields. The proposed PREMOD-AI model introduces a novel hierarchical discriminative AI, which aims to offer rapid and accurate prediction of crop. The method initiates by preprocessing the raw datasets and choosing the more relevant features, improving crop recommendation efficiency while minimizing the time consumption. Next, classification model is employed to analyze both training and testing data based on correlation coefficient for accurately distinguishing the multiple crops. Finally, thePREMOD-AI model utilizes fine-tuning to reduce error rates in multiclass crop classification, thereby increasing the accuracy. A comprehensive assessment was conducted using several performance metrics, including accuracy, precision, recall, F1-score, RMSE, crop recommendation time analysis across different data samples. The quantitative analyzed results demonstrate that the PREMOD-AI model outperforms traditional deep learning methods, achieving higher 95% of accuracy, 95% of precision and 96% of F-score and 97% of recall in crop recommendation and minimizing the 76ms of time. But, the proposed method was not considered for larger datasets, real-time agricultural deployment, and integration with IoT-based smart farming systems. In future work, the proposed method will be employed to apply novel AI and optimization for vast dataset to predict apt crop. Also, the real-world test will be conducted for weather changes and soil differences. In addition, proposed method will be combined with IoT sensors. IoT uses small tools in the field to collect real-time data on the weather and soil.

REFERENCES

[1]J. Madhuri, B. Ramana Reddy, Gauri Kalnoor, M. Indiramma, N. Nagarathna, "Optimizing Crop Recommendations with Improved Deep Belief Networks: A Multimodal Approach", IEEE Access, Volume 13, 2025, Pages 31762 – 31773. **DOI:** 10.1109/ACCESS.2025.3542284

[2] A. Punitha, V. Geetha, "Gorilla troops optimization with deep learning based crop recommendation and yield prediction", International Journal of Cognitive Computing in Engineering, Elsevier, Volume 5, 2024, Pages 494-504. https://doi.org/10.1016/j.ijcce.2024.09.006
[3]Luisa Stracqualursi, "A Lightweight Exponential-Weighted Ensemble for Crop Recommendation", Journal of Agricultural, Biological and Environmental Statistics, Springer, 2025, Pages 1-17.https://doi.org/10.1007/s13253-025-00694-6
[4]Stella Mary Venkateswara & Jayashree Padmanaban, "Interpretable deep learning models for independent fertilizer and crop recommendation", Scientific Reports, volume 15, 2025, Pages 1-24.https://doi.org/10.1038/s41598-025-26910-4
[5]Murali Krishna Senapaty, Abhishek Ray and Neelamadhab Padhy, "A Decision Support System for Crop Recommendation Using Machine Learning Classification Algorithms", Agriculture, Volume 14, Issue 8, 2024, Pages 1-40.https://doi.org/10.3390/agriculture14081256
[6]Hadeeqa Afzal, Madiha Amjad, Ali Raza, Kashif Munir, Santos Gracia Villar, Luis Alonso Dzul Lopez & Imran Ashraf, "Incorporating soil information with machine learning for crop recommendation to improve agricultural output", Scientific Reports, volume 15, 2025, Pages 1-15.https://doi.org/10.1038/s41598-025-88676-z
[7]Muhammad Abu Bakr, Ahmad Jaffar Khan, Sultan Daud Khan, Mohammad Haseeb Zafar, Mohib Ullah and Habib Ullah , "Evaluation of Learning-Based Models for Crop Recommendation in Smart Agriculture", Information, Volume 16, Issue 8, 2025, Pages 1-17.https://doi.org/10.3390/info16080632
[8] A Yaganteeswarudua, Saroj Kumar Biswasa, Varanasi Arunab, Deeksha Tripathi, "Enhancing Transparency in Smart Farming: Local Explanations for Crop Recommendations Using LIME", Procedia Computer Science, Elsevier, Volume 258, 2025, Pages 1993–2005. https://doi.org/10.1016/j.procs.2025.04.450
[9]Md. Sakib Bin Alam, VatcharapornEsichaikul, Aiman Lameesa, Shams Forruque Ahmed, Amir H. Gandomi, "An approach for crop recommendation with uncertainty quantification based on machine learning for sustainable agricultural decision-making",Results in Engineering, Elsevier, Volume 26, 2025, Pages 1-14.https://doi.org/10.1016/j.rineng.2025.105505
[10]Manjit Baishya, Lachit Dutta, "Tiny ML based crop recommendation system for precision agriculture 5.0", Smart Agricultural Technology, Elsevier, Volume 12, 2025, Pages 1-9.https://doi.org/10.1016/j.atech.2025.101247
[11]Amit Bholaand Prabhat Kumar, "ML-CSFR: A Unified Crop Selection and Fertilizer Recommendation Framework Based On Machine Learning", Scalable Computing: Practice and Experience, Volume 25, Issue 5, 2024, Pages 4411-4427. https://doi.org/10.12694/scpe.v25i5.2599
[12]Behnaz Motamedi, Balázs Villányi, "A predictive analytics model with Bayesian-Optimized Ensemble Decision Trees for enhanced crop recommendation", Decision Analytics Journal, Elsevier, Volume 12, 2024, Pages 1-15.https://doi.org/10.1016/j.dajour.2024.100516
[13]Martin Kuradusenge, Eric Hitimana, Kambombo Mtonga, Antoine Gatera, Joseph Habiyaremye, Jackson Ngabonziza, Damien Hanyurwimfura, Placide Rukundo & Angelique Mukasine, "SMART-CYPS: an intelligent internet of things and machine learning powered crop yield prediction system for food security", Discover Internet of Things, Springer, Volume 4, 2024, Pages 1-20.https://doi.org/10.1007/s43926-024-00079-0
[14] Biplob Dey, Jannatul Ferdous, Romel Ahmed, "Machine learning based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables", Heliyon, Elsevier, Volume 10, Issue 3, 2024, Pages 1-13.https://doi.org/10.1016/j.heliyon.2024.e25112
[15]Chetan Raju, Ashoka D.V., Ajay Prakash B.V, "CropCast: Harvesting the future with interfused machine learning and advanced stacking ensemble for precise crop prediction", Kuwait Journal of Science, Elsevier, Volume 51, Issue 1, January 2024, Pages 1-9.https://doi.org/10.1016/j.kjs.2023.11.009
[16]Yashashree Mahale, Nida Khan, Kunal Kulkarni, Shivali Amit Wagle, Preksha Pareek, Ketan Kotecha, Tanupriya Choudhury & Ashutosh Sharma, "Crop recommendation and forecasting system for Maharashtra using machine learning with LSTM: a novel expectation-maximization technique", Discover Sustainability, Springer,Volume 5, 2024, Pages 1-23.https://doi.org/10.1007/s43621-024-00292-5
[17]Mohamed Omar Abdullahi, Abdukadir Dahir Jimale, Yahye Abukar Ahmed & Abdulaziz Yasin Nageye, "Revolutionizing Somali agriculture: harnessing machine learning and IoT for optimal crop recommendations", Discover Applied Sciences, Springer, Volume 6, 2024, Pages 1-11.https://doi.org/10.1007/s42452-024-05739-y
[18]Mohamed Bouni, Badr Hssina , Khadija Douzi and Samira Douzi, "Integrated Iot Approaches for Crop Recommendation and Yield-Prediction Using Machine-Learning", IoT , Volume 5, Issue 4, 2024, Pages 1-16.https://doi.org/10.3390/iot5040028
[19]Md Anisur Rahman, Narayan Ranjan Chakraborty, Abu Sufiun, Sumit Kumar Banshal, Fowzia Rahman Tajnin, "An AIoT-based hydroponic system for crop recommendation and nutrient parameter monitorization", Smart Agricultural Technology, Elsevier, Volume 8, 2024, Pages 1-14.https://doi.org/10.1016/j.atech.2024.100472
[20]SasmitaSubhadarsinee Choudhury, Priya B. Pandharbale, Sachi Nandan Mohanty, Alok Kumar Jagadev, "An Acquisition Based Optimised Crop Recommendation System with Machine Learning Algorithm", EAI Endorsed

Transactions on Scalable Information Systems, Volume 11 Issue 1, 2024, Pages 1-7. https://doi.org/10.4108/eetsis.4003
[21]Sourabh Shastri, Sachin Kumar, Vibhakar Mansotra& Rohit Salgotra, “Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence”, Scientific Reports, volume 15, 2025, Pages 1-16.https://doi.org/10.1038/s41598-025-07003-8
[22]Abid Badshah, Basem Yousef Alkazemi, Fakhrud Din, Kamal Z. Zamli, Muhammad Haris, “Crop Classification and Yield Prediction Using Robust Machine Learning Models for Agricultural Sustainability”, IEEE Access ,Volume 12, 2024, Pages 162799 – 162813. **DOI:** 10.1109/ACCESS.2024.3486653
[23]Farida Siddiqi Prity, MD. Mehadi Hasan, Shakhawat Hossain Saif, Md. Maruf Hossain, Sazzad Hossain Bhuiyan, Md. Ariful Islam & Md Tousif Hasan Lavlu, “Enhancing Agricultural Productivity: A Machine Learning Approach to Crop Recommendations”, Human-Centric Intelligent Systems, Springer, Volume 4, 2024, Pages 497–510.https://doi.org/10.1007/s44230-024-00081-3
[24]Rania A. Ahmed, Walid El-Shafai, Zeinab A. Ahmed, El-Sayed M. El-Rabaie& Fathi E. Abd El-Samie, “High-precision crop recommendation system with stacking ensemble classifiers for optimizing agricultural productivity”, Scientific Reports, Volume 15, 2025, Pages 1-28.https://doi.org/10.1038/s41598-025-09640-5
[25]Hemalatha Gunasekaran, Deepa Kanmani and Ramalakshmi Krishnamoorthi, “Optimized Ensemble Learning for Enhanced Crop Recommendations: Leveraging ML for Smarter Agricultural Decision-Making”, Engineering Proceedings, Volume 82, Issue 1, 2024, Pages 1-8.https://doi.org/10.3390/ecsa-11-20366
[26]Sehrish Munawar Cheema, Ivan Miguel Pires, “AIoT based soil nutrient analysis and recommendation system for crops using machine learning”, Smart Agricultural Technology, Elsevier, Volume 11, 2025, Pages 1-11.https://doi.org/10.1016/j.atech.2025.100924
[27]Gopinath Selvaraj ,Sakthivel Kuppusamy &Menaka Aswathanarayanan, “Sustainable crop recommendation system using seasonally adaptive recursive spectral convolutional neural network for responsible agricultural production”, Geomatics, Natural Hazards and Risk, Volume 16, Issue 1, 2025, Pages 1-23.https://doi.org/10.1080/19475705.2025.2509619
[28]Pedina Sasi Kiran, Gembali Abhinaya, Smaraneeka Sruti and Neelamadhab Padhy, “A Machine Learning-Enabled System for Crop Recommendation”, Engineering Proceedings, Volume 67, Issue 1, 2024, Pages 1-11. https://doi.org/10.3390/engproc2024067051
[29] Gurpreet Singh & Sandeep Sharma, “Enhancing precision agriculture through cloud based transformative crop recommendation model”, Scientific Reports, volume 15, 2025, Pages 1-22.https://doi.org/10.1038/s41598-025-93417-3
[30]Anitha Palakshappa, Sowmya Kyathanahalli Nanjappa, Punitha Mahadevappa, Sinchana, “Smart irrigation with crop recommendation using machine learning approach”, Bulletin of Electrical Engineering and Informatics, Volume 13, Issue 3, 2024, Pages 1952-1960. https://doi.org/10.11591/eei.v13i3.6103
[31] Mohanappriya K., Vennila C. & Keerthivasan N., “AgroDualNet: a dual deep learning-based crop disease forecasting and fruit ripening detection”, Scientific Reports, Springer, Volume 16, April 2026, Pages 1-10. . https://doi.org/10.1038/ s41598-026-49002-3
[32] Thandra Jithendra, A Divya, Ibrahim Aljubayri, T Haritha and Mohammad Zubair Khan, “Two-stage rainfall forecasting and crop classification using puma-optimized ANFIS–SVM”, Smart Agricultural Technology, Elsevier, Volume 13, March 2026, Pages 1-21. https://doi.org/10.1016/j.atech.2025.101719
[33] S. Poornima, Prafful Mishra, Rohan R. Mahishi, Deepak Chand Chaganti, Pushpalatha M & Raghavendra B. Jana, “Optimizing crop selection for sustainable agriculture: a compound ensemble approach integrating machine learning and IoT-based sensors”, Scientific Reports, Volume 16, Issue 1, March 2026, Pages 1-21. doi: 10.1038/s41598-026-40772-4.