Proceedings of the 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC-2024)
IEEE Xplore Part Number: CFP24BC3-ART; ISBN: 979-8-3503-7519-0

Crop Recommendation System using Antlion Optimization and Decision Tree Algorithm

Dr. J. Avanija
Professor, AI&ML,
School of Computing,
Mohan Babu University
Sree Sainath Nagar,Tirupati,
Andhra Pradesh, India.
avans75@yahoo.co.in

Keerthi Ambati
UG Scholar,Department of CSE,
Sree Vidyanikethan Engineering College
Sree Sainath Nagar,Tirupati,
Andhra Pradesh, India
keerthiambati18@gmail.com

Likitheswari Naraganti,
UG Scholar, Department of CSE,
Sree Vidyanikethan Engineering College
Sree Sainath Nagar,Tirupati,
Andhra Pradesh, India
naragantilikitha@gmail.com

Sai Sahith Derangula
UG Scholar,Department of CSE,
Sree Vidyanikethan Engineering College
Sree Sainath Nagar,Tirupati,
Andhra Pradesh, India
saisahith03@gmail.com

Tanujasree Nashina,
UG Scholar, Department of CSE,
Sree Vidyanikethan Engineering College
Sree Sainath Nagar,Tirupati,
Andhra Pradesh, India
sree200218@gmail.com

Abstract— **Agriculture in India is a cornerstone of employment and economic growth, contributing significantly to the nation's GDP. However, conventional farming practices often lead to suboptimal crop yields and financial losses due to inadequate knowledge of crop selection based on soil types and climate variability. To address this challenge, an advanced crop recommendation system is developed, integrating Ant Lion Optimization and Decision Trees (ALO-DT) to suggest optimal crop types based on ecological factors. The proposed model, trained on a Kaggle dataset, combines global optimization with decision-making precision to automatically recommend crops suited to diverse agricultural conditions. By synergizing optimization techniques with data-driven precision, the system provides tailored recommendations for varying ecological contexts, empowering farmers to make informed decisions. These recommendations not only enhance agricultural performance but also support sustainable farming practices, fostering food security and economic resilience. The proposed crop recommendation system serves as a valuable tool in agricultural decision-making, offering guidance to farmers and contributing to the advancement of sustainable agribusiness practices in india.**

Keywords— Nature Inspired Optimization, Agriculture, Crop Recommender System, Ant Lion Optimization(ALO), Ant Lion Optimization-Decision Tree(ALO-DT), Crop Yield, Humidity.

I. Introduction

In a developing country like India farming plays a major role by providing food security and a large number of employment. Most likely 58% of India's population, a huge sum of 600 million people, makes their livelihood form agriculture whether directly or indirectly. The agriculture sector provides a massive 14% in the India’s GDP, which shows the power it holds. Agriculture is not just about food security, there exists many things beyond it like agriculture promotes economic backbone, international connections, social and cultural fabric, trade, and global commerce. Over the years, farmers have widely depended on their previous knowledge gained from their ancestors and their own experience to choose the best crops for their field. However, with the increasing complexity of agricultural systems and drastically varying environmental issues, the data-driven methods are widely justified to be important. To overcome these situations and to catch up with the modern needs, promoted the use of Crop Recommendation System. A Crop Recommendation System is a practical development and a creative extension of technology in agriculture that utilizes data-driven insights, machine learning algorithms, and agronomic knowledge to give farmers a personalized and optimized recommendations about the which crop to be selected, the type of cultivation practices based on availability of water, and managing resources efficiently. This model targets to improve the agriculture productivity, to maximize the efficiency of resources used, and then to ultimately improve the overall profitability and sustainability of farming operations. The proposed system takes various factors into consideration like type of soil, climatic conditions, availability of water, marketing demand, and the farmer's preferences. One such significant model of approaching crop recommendation systems is the Ant Lion Optimization (ALO). ALO is a nature inspired optimization algorithm which is developed depends on the hunting behavior of antlions. This method copies the process of an antlion trapping unsuspecting ants in its conical pit. ALO is used to predict the suitable combination of crop types and planting patterns that generates maximum yield and profits while considering environmental constraints like rainfall, nitrogen%, potassium%, phosphorous%, PH in the aspect of crop recommendation.

II. Related Works

A lot of work that has previously been done indicates that the machine learning models and nature-inspired algorithms alone are not efficient for crop recommendation which must result in less accurate results. so proposed methodology used hybrid techniques for the crop recommendation system i.e. antlion optimization along with the decision tree algorithm, the main aim of ALO optimize the hyperparameters used for the decision tree for enhancing the crop recommendation systems.

[1] introduces a rule-based crop management recommendation system, successful in a Chiang Mai greenhouse but potentially limited by diverse soil types. Future enhancements could involve machine learning adaptation and collaboration with local experts. The study in [2] describes a machine learning-based crop prediction system using soil sensors to analyze temperature, moisture, and nutrient data. The system suggests suitable crops and employs a program to assess plant health, aiming to aid farmers in decision-making. Challenges exist in adapting the system to different soil types, with ongoing research on rule development for improved effectiveness at scale.

Paper [3] introduces a hybrid system using Ant Colony Optimization and Deep Learning for accurate crop yield predictions in precision agriculture. Despite satisfactory results, challenges remain in ensuring precision agriculture systems provide accurate recommendations to avoid potential losses. [4] integrates the Particle Swarm Optimization-Multi-Layer Deep Neural Network model for crop recommendation, boasting a 95.49% accuracy surpassing other methods. It suggests hybrid improvements while acknowledging challenges like overfitting with limited data and scalability issues in costly machine-learning techniques for local datasets. [5] introduces a highly precise crop recommendation system using machine learning that uses ensemble techniques with great average accuracy in classifying soil for optimal Kharif and Rabi crop selection, emphasizing precision agriculture's role in boosting yield. The research study conducted in [6] introduces AgroConsultant, a knowledge driven model that recommends crops designed to aid farmers in decision-making. By analyzing factors like sowing season, geographical location, soil, temperature, and rainfall, the system employs machine learning algorithms and big data analytics to offer precise crop recommendations, aiming to enhance farmers' decision accuracy. The research carried out in [7] presents a crop recommendation system that uses algorithms in machine learning to predict profitable crops to farmers depending on factors like rainfall and soil type. Implemented as a user-friendly mobile app, the system achieves a promising 95% accuracy in crop yield prediction, also providing recommendations on optimal fertilizer usage timing. [8] utilized deep learning and IoT for precise agricultural information classification but faces hurdles like the need for extensive labeled data and scalability verification. Restrictions on training datasets impact model generalizability, and the study talks about a trade-off between classification speed and accuracy during sample set pruning, enhancing the importance of mentioning these issues for real life applications. [9] showcases the overall review of intelligent recommender systems in agriculture, prioritizing the information knowledge. It includes content analysis to categorize research, covering concepts, technologies, and applications. The recognized drawbacks encompass data sparsity, cold start issues, and the imperative to increase the accuracy of recommendation, highlighting challenges in the agriculture sector. [10] proposes a new method of predicting the yield rates associated with crops and detect how agro products prices should vary is presented. The paper introduces an ELSTM model-based on state of the art computationally intensive data driven approaches, with which a yield prediction accuracy of 85% is achieved.

In [11], the reviews about using big data and AI in precision agriculture to enhance the productivity. It standouts the applications and challenges, including data collecting issues, resource limitations for small farms, and ethical concerns about data accessibility and cost. Managing data volume and ensuring real-time scalability are identified as significant challenges in the precision agriculture landscape. In [12] Madhuri and Indiramma's research introduces an innovative crop recommendation system using Artificial Neural Networks, incorporating soil, crop characteristics, and climate parameters to recommend maize, Finger millet, Rice, and sugarcane based on location-specific data from Karnataka, India. Their model achieved an impressive 96% accuracy, highlighting its potential for advancing sustainable agriculture through tailored recommendations.

The study in [13] presents an machine learning technique based on ensembling method for predicting crops, addressing data limitations and proposing future work on deep learning methods and market price analysis. It references related works on crop production prediction, highlights the superiority of the Kernel Ridge Regression model, and emphasizes the importance of data availability, funding sources, and author contributions for transparency and future research directions. The research carried out in [14] introduces a three-tiered framework for recommending 22 crops in agriculture, employing correlation and distribution analyses, along with ensembling techniques like random forest, decision trees and support vector machines. Naïve Bayes achieves the highest accuracy at 99.54%, while challenges and future research directions are discussed. The study in [15] addresses the importance of scientifically proven crop recommendation methods for enhancing agricultural productivity and profitability in developing countries. This paper explores different machine learning algorithms that is gradient boosting classifier, KNN, and random forest classifier, followed by an ensemble approach using a voting classifier for improved outcomes. Significantly, it discusses the results of these algorithms based on metrics like accuracy, execution time, logistic regression and f1 score along with the utilization of cloud computing for efficient processing, user interaction and tabulates the comparisons.

III. Proposed Methodology

The technique utilizes both the AntLion optimization and the Deep Learning concepts, which in return promotes a hybrid approach to improvise the model accuracy for crop recommendation. The proposed method comprises of multiple stages, containing optimization, evaluation, and visualization, resulting in the performance improvement of the decision tree and the more precise crop recommendations. The main reason behind selecting the ant lion optimization is that it is used to find the global optimum in search space efficiently, it explores the hyperparameter space thoroughly to find the combination that yields the best performance.

1. Dataset:

This dataset can be used as input for the precision agriculture, which helps to suggest appropriate crop and methods to the farmers considering their needs. The dataset is made from years long of collected information on datasets of rainfall, climate, and fertilizers used in the field specifically for India. This dataset consists of 2201 records with the following data fields:

NPK Ratio: The ratio of Potassium, Nitrogen and Phosphorous availability in the field
Temperature: Temperature in the calculated in Celsius degrees.
Humidity: humidity of the field in percentage.
pH: Soil PH in the field
Rainfall: Measured Rainfall in mm

The above datafield values helps to generate best accurate models to recommend suitable crop for the filed. The dataset plays an important role in the precision agriculture. The above utilized dataset is find using the link https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset/data.

2. Preprocessing:

In this model, the dataset which is collected conducts various pre-processing procedures such as data cleaning; data combinatory management for integration to form a 1) single coherent record; performs transformations that 2) accommodate storage and retrieval from optimal databases through a capacity of ours listed parameters in Python. Then the given dataset is split into a training and test set thereafter stored separately for future applications.

3. Model building:

The inputs of our model are the crop data from different resources such as kaggle. The collected input data is pre-processed by performing the steps such as integrating the data, cleaning the data and data reduction. The information is then passed into the Antlion optimization algorithm which optimizes the parameters for the decision tree algorithm as shown in Fig.1. The data is split into the testing and training datasets, the data that is trained passed into the Machine-learning model that is our decision tree and trained and then we test the model based on our testing values. Finally, we can predict the crop using the decision tree algorithm and then we can evaluate our performance metrics.

4. Architecture:

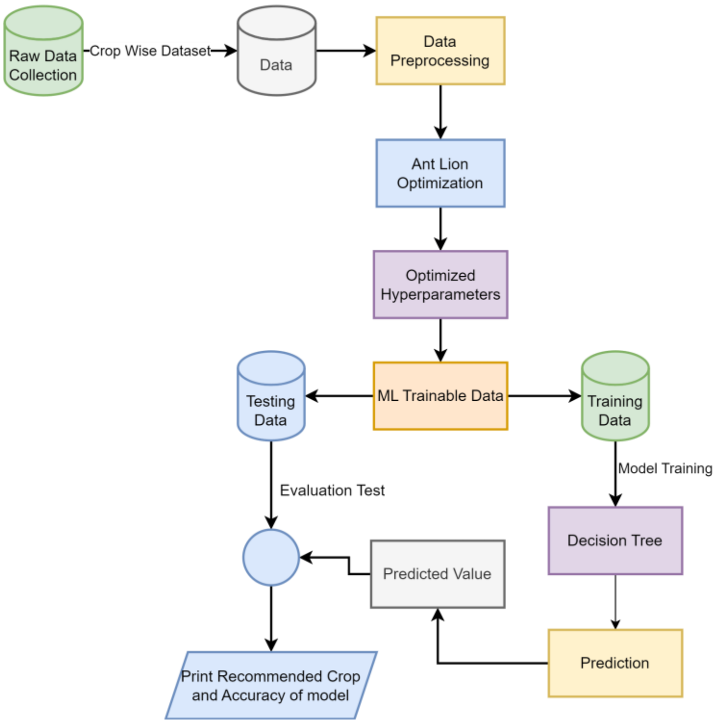

Fig.1 . Architecure of ALO-DT

5. ALGORITHM:

Input:
"Crop Recommendation" Dataset is taken from Kaggle.
The hyperparameters used are:

- Lower bound values and upper bound values for depth, random state and criterion
- Number of ant lions
- Maximum number of iterations
- Problem size (In this case it is 3- depth, random state and criterion)

Output:
Suitable crop type that can be grown on that land.
Algorithm:
Step 1: Load the crop Recommendation dataset.
Step 2: Split the dataset into features(x) and target variable(y) and next split the data into training and testing datasets.
Step 3: Initialize the population of ant lions.

- Each ant lion represents a potential solution to the optimization problem.
- Each ant lion has 3 parameters – depth, random state, criterion.
- Randomly initialize the population of ant lions within the search space which is between the lower bounds and upper bounds.

Step 4: Evaluate the fitness of each ant lion using objective function. Here the objective function is decision tree optimization. Fitness calculation is represented in equation (1)

$$\text{Fitness} = \text{Accuracy} = \frac{\textit{Number of Correct Predictions}}{\textit{Total Number of Predictions}} \quad (1)$$

Step 5: Sort the ant lions(solutions) based on their fitness values in descending order.
Step 6: Retain the top half of solutions as elite ant lions.
Step 7: Generate new solutions or new ant lions randomly within the search space defined between lower bound and upper bound for the non-elite group.
Step 8: Combine elite and non-elite ant lions to form the next generation.
Step 9: Repeat steps 4 to 8 for a predefined number of iterations.
Step 10: Select the best solution found by Ant Lion Optimization.
Step 11: Construct the Decision tree using the best hyperparameters found- max depth, random state and criterion. Train the decision tree on training dataset.
Gini impurity used to measure the impurity of samples is represented in equation (2):

$$Gini(P)=1 - \sum_{i=1}^{n} P_i^2 \quad (2)$$

Entropy is used to measures the impurity or uncertainty of samples represented in Equation (3):

$$Entropy(P)= - \sum_{i=1}^{n} P_i^2 \, log_2(P_i) \quad (3)$$

Step 12: Now, Evaluate the performance of the best decision tree model on the test set.
Step 13: Give input parameters of the land like (Content of Nitrogen, Phosphorous, Potassium etc. in soil), the model will recommend the best crop that can be grown on that land.
Step 14: Output: Optimal hyperparameters derived from Ant Lion Optimization and Prediction of suitable Crop type for the specified land.

IV.Performance Analysis

1.Accuracy:

Accuracy measures the overall effectiveness of crop recommendations made based on hybrid model. It is calculated by replacing the incorrect instances in numerator with True Positives and True Negatives.it is calculated as:

2. PRECISION:

It is thus the measurement of predictive accuracy by hybrid model for positive crop predictions. The ratio of correctly determines its predicted positive instances to the total positive predictions.it is calculated as: As such, if a service delivered by one party is not fit for use or fulfilling the purpose it was intended to serve in terms of products and related conditions supplied under laws demanding publicity prior to their usage.

3. RECALL:

Recall measures the hybrid model's capability to correctly retrieve all true positives amongst the crops recommended. It is derived as the quotient of True Positives and total number if actual positives i.e., sums for True Positive and False Negatives values (TPT/ TPT+FN).

4.F1-SCORE:

Harmonic mean of recall and precision can be calculated using F1 score. It provides a very well balanced results of the hybrid optimization model's performance.

5. ERROR RATE:

The metric error rate quantifies the proportion of incorrectly classified crop recommendations in the hybrid model. It is determined by the ratio of mismatched identified instances i.e.; the False Negatives and False Positives to the total number of instances.

TABLE I: PERFORMANCE ANALYSIS

Methods / Metrics	Decision Tree	KNN	R-Forest	Neu-net	ALO-DT
Accuracy	90.47	88.78	91.65	92.76	98.86
Precision	85.90	81.27	85.16	80.11	98.86
Recall	89.38	89.37	90.89	91.38	98.86
FMeasure	87.60	85.12	87.93	85.37	98.86

V.Results & Discussion

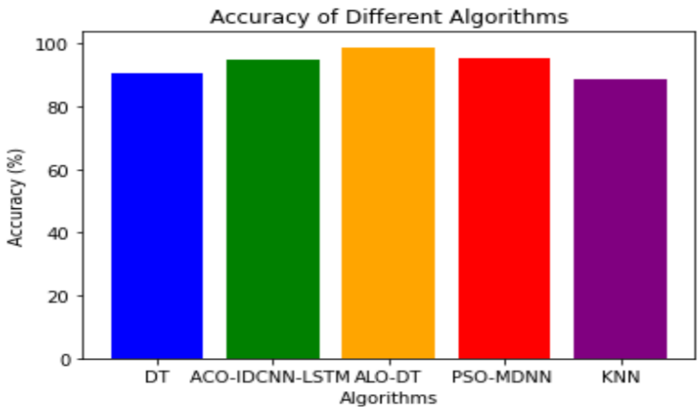

Fig.2. Performance Comparison based on Accuracy

Results obtained from the hybrid model the antlion optimization along with the decision tree algorithm shows that ALO-DT model has high accuracy than the other compared models. It is shown in Fig.2. Due to hyperparameter optimization using ALO by exploring search space and constructing decision tree on those optimized parameters helps to achieve the high accuracy. After constructing the decision tree , Train the model using the training data and use it to make predictions on unseen data by providing input parameters(e.g. Soil PH, temperature, rainfall etc..) for the prediction and model will output the recommended crop based on its learned patterns and decision rules. The predicted results of the proposed hybrid model are more accurate than that of other models. This means it is more accurate in predicting the correct crop. This combination of antlion and decision tree has provided better results.

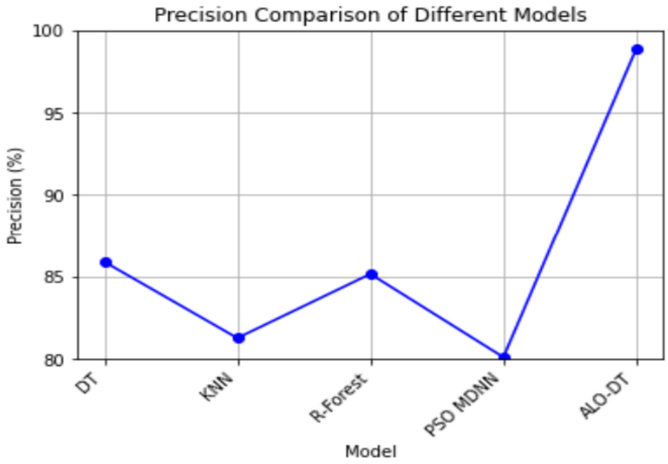

Fig. 3. Performance Comparison based on Precision

Fig.3 shows that ALO-DT(Ant Lion Optimization-Decision Tree) model has more precision value when compared to other models. Fig.4 shows the training loss and validation loss Of the hybrid model(ALO-DT) over each epoch. Fig.5 shows the training accuracy and testing accuracy of the hybrid model(ALO-DT) over each epoch

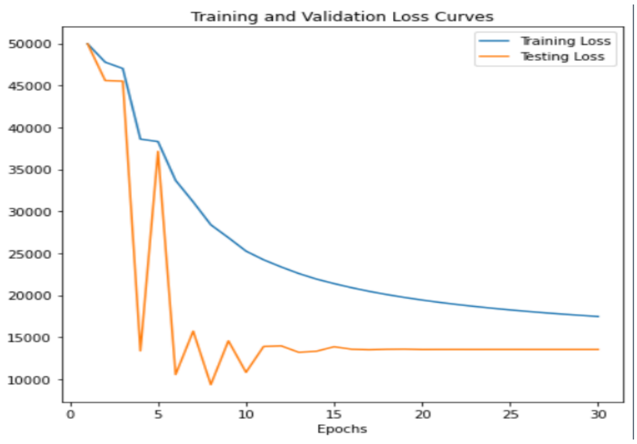

Fig .4. Training and Validation Loss Curves of ALO-DT

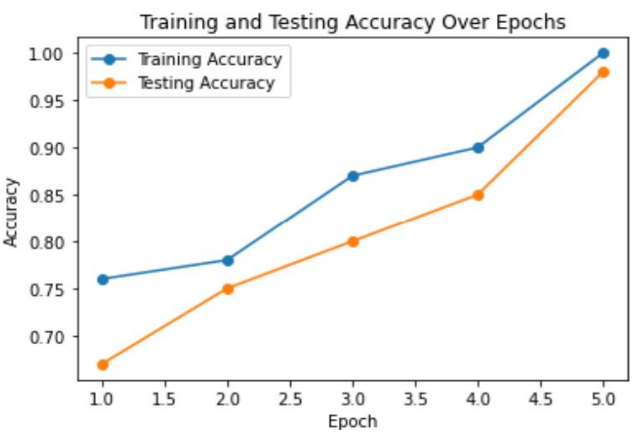

Fig.5. Training and Testing Accuracy of ALO-DT

VI. CONCLUSION

This study has presented a novel hybrid model for crop recommendation in precision agriculture, integrating Antlion Optimization with Decision Tree methodologies. Our approach leverages the foraging behavior of antlions to optimize parameters for Decision trees, contributing to improved accuracy and robustness in predicting optimal crop choices.

The experimental results demonstrate the efficiency of the proposed hybrid model, showcasing superior performance compared to traditional Decision Tree and Antlion Optimization methods. Notably, our model achieved the highest accuracy of 98.86% in recommending crops for diverse agricultural scenarios. In conclusion, the hybrid model demonstrates significant potential in enhancing crop selection processes, contributing to sustainable and efficient agricultural practices. The Future scope of this project involves integration of real time weather and soil data will emphasize the accuracy and relevance of crop recommendation and Combined utilization of Geographic Information System (GIS) technology enables region specific customization of suggestions by taking into account the local soil and climate data. We believe the findings shown in this paper leads to many upcoming improvements in precision farming, where the combination of optimization algorithms with machine learning proves instrumental in addressing complex decision-making challenges for farmers.

REFERENCES

[1] Paweena Suebsombut, Aicha Seklouli-Sekhari, Pradorn Sureephong, Abdelaziz Bouras. “Rule-based recommendation system to support crop lifecycle management” The Journal of Modern Project Management, 2021, 9 (2), ⟨10.19255/JMPM02717⟩.

[2] D. Gosai, C. Raval, R. Nayak, H. Jayswal, and A. Patel, “Crop Recommendation System using Machine Learning,” *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, pp. 558–569, Jun. 2021.

[3] K. Mythili and R Rangaraj, “Crop Recommendation for Better Crop Yield for Precision Agriculture Using Ant Colony Optimization with Deep Learning Method,” *Annals of the Romanian Society for Cell Biology*, pp. 4783–4794, Apr. 2021.

[4] K. Mythili and R. Rangaraj, “Deep Learning with Particle Swarm Based Hyper Parameter Tuning Based Crop Recommendation for Better Crop Yield for Precision Agriculture,” *Indian Journal of Science and Technology*, vol. 14, no. 17, pp. 1325–1337, May 2021.

[5] A. Sharma, A. Jain, P. Gupta, and V. Chowdary, “Machine Learning Applications for Precision Agriculture: A Comprehensive Review,” *IEEE Access*, vol. 9, pp. 4843–4873,2021,doi:https://doi.org/10.1109/access.2020.3048415.

[6] Z. Doshi, S. Nadkarni, R. Agrawal, and N. Shah, “AgroConsultant: Intelligent Crop Recommendation System Using Machine Learning Algorithms,” *IEEE Xplore*, Aug. 01,2018. https://ieeexplore.ieee.org/abstract/document/8697349 (accessed Oct. 09, 2022).

[7] S. M. PANDE, P. K. RAMESH, A. ANMOL, B. R. AISHWARYA, K. ROHILLA, and K. SHAURYA, “Crop Recommender System Using Machine Learning Approach,” *2021 5th International Conference on*

Computing Methodologies and Communication (ICCMC), Apr. 2021, doi: https://doi.org/10.1109/iccmc51019.2021.9418351.

[8] H. Yang, T. Wen, X. Yang, and H. Lin, "Deep Learning Agricultural Information Classification Combined With the Internet of Things Technology in Agricultural Production and Economic Management," *IEEE Access*, vol. 10, pp. 54713–54719, 2022, doi: https://doi.org/10.1109/access.2022.3176643.

[9] C. Song and H. Dong, "Application of Intelligent Recommendation for Agricultural Information: A Systematic Literature Review," *IEEE Access*, vol. 9, pp. 153616–153632, 2021, doi: https://doi.org/10.1109/access.2021.3127201.

[10] O. N. Ovelade and A. E. Ezugwu, "Ebola Optimization Search Algorithm: A new nature-inspired metaheuristic algorithm for global optimization problems," *2021 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, Dec. 2021, doi: https://doi.org/10.1109/icecet52533.2021.9698813.

[11] "(PDF) Big Data and AI Revolution in Precision Agriculture: Survey and Challenges," *ResearchGate*, doi: https://doi.org/10.1109/ACCESS.2021.3102227.

.

[12] J. Madhuri and M. Indiramma, "Artificial Neural Networks Based Integrated Crop Recommendation System Using Soil and Climatic Parameters," *Indian Journal of Science and Technology*, vol. 14, no. 19, pp. 1587–1597, May 2021, doi: https://doi.org/10.17485/ijst/v14i19.64.

[13] M. Hasan *et al.*, "Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop cultivation," *Frontiers in Plant Science*, vol.14,aug2023,doi:https://doi.org/10.3389/fpls.2023.1234555.

[14] R. K. Ray, S. K. Das, and S. Chakravarty, "Smart Crop Recommender System-A Machine Learning Approach," *IEEEXplore*,Jan.01,2022. https://ieeexplore.ieee.org/document/9734173(accessed Apr. 18, 2022).

[15] Gourab Dhabal, Jaykumar Lachure, and Rajesh Doriya, "Crop Recommendation System with Cloud Computing," *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, Sep.2021,doi:https://doi.org/10.1109/icirca51532.2021.9544524.