ISSN: 2182-2069 / E-ISSN: 2182-2077

Optimizing Sustainability of Smart Farming with IoT-Assisted Novel Framework

Ashvin B Amale[1*], Dr.K.V. Pradeep[2], A.J. Rishabh Bhardwaj[3], Dr. Neeta Mazumdar[4], Dr. Sachin S Pund[5], and Bhuvana Jayabalan[6]

[1*]Assistant Professor, Department of Mechanical Engineering, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India. ashvin.amale@gmail.com, https://orcid.org/0009-0003-5653-5231

[2]Associate Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Tamil Nadu, India. chennaipradeep.kv@vit.ac.in, https://orcid.org/0000-0003-2094-5450

[3]Centre of Research Impact and Outcome, Chitkara University, Rajpura, Punjab, India. rishabh.bhardwaj.orp@chitkara.edu.in, https://orcid.org/0009-0009-6075-8837

[4]Associate Professor, Department of Mathematics, Government College of Arts, Science and Commerce, Quepem, Goa, India. neetamazumdar@gmail.com, https://orcid.org/0000-0002-8971-1065

[5]Assistant Professor, Department of Mechanical Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur, Maharashtra, India. pundss@rknec.edu, https://orcid.org/0000-0002-5616-2469

[6]Associate Professor, Department of Computer Science and Information Technology, Jain (Deemed to be University), Bangalore, Karnataka, India. j.bhuvana@jainuniversity.ac.in, https://orcid.org/0000-0002-8372-6311

Received: September 04, 2025; Revised: October 11, 2025; Accepted: December 04, 2025; Published: February 27, 2026

Abstract

An essential component of the Indian economy is agriculture. Precision farming, another call for smart farming, is a novel method of farming that incorporates modern innovations to improve production, sustainability, and performance over traditional techniques. These involve strong data security measurements, the reliance on high-quality sensor data, and possible integration issues with current farming systems. This study proposed a methodology for smart farming using IoT sensor statistics incorporated with a unique Dynamic Artificial Fish Swarm Driven XGBoost (DAFS-XGBoost) model. Initially, research was conducted on a Kaggle dataset. Substantially, the "Crop recommender CSV" file, containing vital IoT sensor records, which include temperature, humidity, rainfall, soil potential of Hydrogen (pH), and nutrient ratios, changed into utilized. Data cleaning procedures were implemented to rectify inaccuracies, incompleteness, and redundancy. The dataset has been split into subsets for testing and training to improve the model. The proposed DAFS-XGBoost approach, carried out with the use of MATLAB, is designed to enhance predictive accuracy while maintaining low latency and efficient hardware usage. The findings demonstrated that the proposed DAFS-XGBoost method achieves effective recall, accuracy, precision, f1-score, training time, and MSE rates as 93.76%, 96.73%, 94.61%, 90.12, 0.0006, and 0.19. The framework's potential to combine outstanding performance and a hardware-friendly design to promote the long-term viability of smart farming is confirmed by simulations.

Journal of Internet Services and Information Security (JISIS), volume: 16, number: 1 (February-2026), pp. 31-48.
DOI: *10.58346/JISIS.2026.I1.003*

*Corresponding author: Assistant Professor, Department of Mechanical Engineering, Yeshwantrao Chavan College of Engineering, Nagpur, Maharashtra, India.

Keywords: Smart Farming (SF), Agricultural, Internet of Things (IoT), Crop Prediction, Dynamic Artificial Fish Swarm Driven XGBoost (DAFS-XGBoost).

1 Introduction

One of the major sectors in the economy of India is agriculture; this is because it provides a large percentage of employment and food. However, it is now faced with many challenges that challenge its stability and development. Some of these are population growth especially when it is growing at a rather alarming rate per year which puts pressure as far as food demand is concerned, climate change which is rather unpredictable and has adverse effects as far as farming is concerned, and labor shortage which may hinder effective management of the farms and productivity (Javaid et al., 2022). Such problems are quite modern, yet conventional agriculture practices that have always worked before cannot handle them. This has brought into sharp focus the necessity to come up with efficacious approaches to enhance the sector's future sustainability (Mohamed et al., 2021; Charania & Li, 2020)

In this connection, some new useful tools are appearing in agriculture, one of which is precision agriculture or smart farming. Therefore, this approach incorporates the use of different high technologies to assist in agricultural practices with the main goals of increasing production, efficiency, as well as sustainability (Lytos et al., 2020). Precision agriculture is thus a contemporary approach to farming that employs advanced technologies like cloud computing, artificial intelligence, and the Internet of Things. Through the integration of these technologies, smart farming empowers farmers with better decision-making and more efficient handling of the resources available to them (Mahbub, 2020; Dhanaraju et al., 2022).

Cloud computing is an effective tool for handling large volumes of agricultural data and analyzing it in real time. Such data is analyzed by the AI algorithms to present farming solutions that can be looked at in an anticipatory nature to enable farmers to look forward to some difficulties and even plan their strategy (Patil & Khairnar, 2021). Sensors and actuators, which are components of IoT devices, are vital in smart farming by providing regular data on environmental factors (Nambiar & Varma, 2023). These objectives can be attained through the use of this data to measure the health of the soil and predict water stress, and disease detection among the crops required for high productivity and sustainability of the crops (Idoje et al., 2021).

However, despite such developments, there are considerable blind spots that need to be addressed to realize the full potential of smart farming networks from a sustainability standpoint. Current use cases of IoT in agriculture have emerged with issues related to do with precise data acquisition, real-time sensing, and management of Big Data (Novak & Vacek, 2023). These limitations decrease the possibility of success of smart farming, which is aimed at increasing yields and improving sustainability (Monica Nandini, 2024). This way, the presented framework seeks to address the current problems that affect sustainable methods of agriculture and provide top benefits when it comes to resource utilization and stability of long-term agricultural activities.

Khan et al., (2022) enhanced the system's energy-efficiency capacity by implementing a novel mix of intelligent smart irrigation technologies. Additionally, the HSSC method was used to construct and choose the best head clusters. Considering the new model to the standard performance, the findings of the suggested method indicated that it consumed significantly less energy and provided a longer network lifespan. Pal et al., (2023) proposed an ICISF-DDO system that improves the lifespan and cost-effectiveness of automated agricultural infrastructures. The suggested scheme's conceptual structure and statistics development approach were effectively provided. Mahajan et al. (2021) examined the clustering techniques that were given, which were not practical and resulted in increased latency and delay for different SF applications. The specifications for SF systems for remote farm surveillance and decisions in rural areas were the focus of the development of the CL-IoT protocol, an efficient and flexible standard.

Anand et al., (2021) offered ASN, a DL system that used multi-scale focus semantic division of UAV-obtained pictures to automatically detect farmland abnormalities. To improve the effectiveness of agricultural precision methods, the suggested model was helpful for crop and farmland management (Neethu & Ramyaprabha, 2025). The effectiveness of contemporary agricultural methods can be improved by using IoT-enabled smart technologies to monitor crops and farmland (Kumar, 2024). Remote sensing using UAV was an effective method for gathering agricultural-related picture data. Singh and Singh (2023) presented an approach based on fog with extra IoT assistance, which showed promise for globally accessible smart agriculture architecture (Bahmaid & Ghaleb, 2024). The findings suggested that the literature search on the Web of Science database feature should be utilized to enhance the IoT-Fog instructional materials for the creation of smart agricultural goods.

Table 1: Depicts the existing works on the sustainability of SF with IoT

Reference	Technology Used	Findings / Use Case	Benefits	Drawbacks
Khan et al. (2022)	Intelligent smart irrigation technologies, HSSC method	Significant energy efficiency improvement	Energy consumption reduction	Lacks real-world validation
Pal et al. (2023)	ICISF-DDO system	Automated Farming’s Viability	Enhanced lifespan	Limited validation
Mahajan et al. (2021)	CL-IoT protocol, Cross-layer CH selection	Reduced latency, and scalability for SF applications	Reduced latency and improved scalability	Increased latency and costly implementation
Anand et al. (2021)	ASN, Multiscale attention semantic segmentation	Automated detection of farmland abnormalities	Improved precision agriculture	Complex implementation
Singh and Singh (2023)	Fog-based model, IoT assistance	Globally accessible smart agriculture architecture	Globally accessible architecture	Overstated benefits
Sukanya and Ramachandram (2023)	WSN, Edge, and Cloud-based computing	Effective soil data collection, processing, and storage	Effective data gathering, processing, and storing	High complexity
Alrowais et al. (2022)	HLBODL-WDSA model, IoT devices	Plant detection using IoT devices	Cloud-based analysis	Latency issues
Gupta and Nahar (2023)	Cross-layer-based clustering and routing model	Long-distance data transfer for smart farming	Long-distance data transfer	Complexity overload
Dahane et al. (2022)	Edge-IoT Cloud platform, DL approach	Monitoring and forecasting agricultural water demand	Monitoring and forecasting	Connectivity limitations

Sukanya and Ramachandram (2023) presented a three-tiered architecture for gathering, processing, and storing valuable soil data from farming regions. It depended on a combination of Based on the cloud and edge-based WSN approaches. To gather data on various soil component types, the sensors were dispersed randomly throughout the network region. To facilitate effective communication, the sensors were aggregated according to distance utilizing the K-means clustering technique, which was developed on Levy flight. The new HLBODL-WDSA model was presented by (Alrowais et al., 2022). The suggested HLBODL-WDSA model’s main objective was to detect the plants in the images by employing IoT devices to capture plants. Initially, the farm images were collected by the IoT devices according to the HLBODL-WDSA model, which then sends images to the remote server for analysis.

For WSN-assisted IoT networks for SF, Gupta and Nahar (2023) suggested a method of cross-layer-driven routing and clustering model that enabled long-distance data transfer that was both expandable and cost-effective. The system was initially divided into several clusters using the fuzzy k-medoids clustering methods, as the creation of clusters has a considerable effect on power usage. Utilizing the NS2, the suggested protocol was simulated. The simulation outcomes were evaluated regarding the end-to-end latency, communication overhead, transmission expense, average energy usage, and duration of the network. Dahane et al. (2022) suggested an Edge-IoT Cloud platform for monitoring and forecasting farmers' capacity for satisfying agricultural water demands in the event of minimal rainfall, utilizing a DL approach. Based on the initial findings, they suggested their approach provided a solid place to start when it comes to enabling small-scale farmers to make superior choices through low-cost SF. Table 1 shows that the summary of related works.

Contribution and Motivation of the Work

The motivation and contribution of the work are discussed in this part. IoT-enabled SF has advantages and disadvantages. Adoption may be confined via the scale of the initial expenditure required, particularly for smaller farms. Major boundaries to implementation encompass technical complexity and the requirement for specialized understanding. Operations in remote places with restricted rights of entry can be disrupted with the aid of reliance on internet connectivity. Operational needs are accelerated using the strength desired for everyday protection and continuous operation. To overcome these limitations and ensure the sustained implementation of IoT-enabled SF strategies is critical. These limitations motivated me to develop a novel method called Dynamic Artificial fish swarm driven XGBoost (DAFS-XGBoost) for the sustainability of SF with IoT. Therefore, the following is a summary of this work's contributions:

Data Collection from Kaggle: Obtained the "IoT crop and soil" dataset from Kaggle, particularly making use of the "Crop recommender. CSV" record, which incorporates essential IoT sensor facts, inclusive of temperature, humidity, rainfall, soil pH, and nutrient ratios crucial for crop prediction.

Data Cleaning: It is the process of altering or eliminating data that is lacking, structured incorrectly, inaccurate, repeated, or unnecessary to make it accessible to study.

Data Splitting: The initial stage of creating a model involves creating testing and training sets from the dataset, with training data as the testing set and the initial set as the finalized model fit.

Model Building with Dynamic Artificial Fish Swarm Driven XGBoost: Developed a suggested technique named "Dynamic Artificial Fish Swarm Driven XGBoost" for crop identity, leveraging system learning class algorithms. This innovative approach combines the strength of XGBoost with dynamic synthetic fish swarm optimization to ensure the accuracy and efficiency of crop prediction.

This work is structured as follows: The methods and materials are described in part 2. Part 3 presents the model assessment and part 4 concludes the study.

2 Proposed Methodology

i. Data Collection: Acquire information about the computer hardware system, such as performance measurements, log files, and sensor readings. Data was collected via Kaggle. The IoT farming crops and soils dataset was utilized.

ii. Data Cleaning: It involves deleting or altering data that is formatted incorrectly, lacking, redundant, inaccurate, or unnecessary to make it available for analysis.

iii. Splitting the Data: Creating testing and training groups from the dataset is the first phase in creating the approach. Training data is the first set of data needed to train machine learning (ML) approaches. A set of tests is a collection of information that is utilized to accurately evaluate whether a finalized model fit is applied to the training sample.

iv. Model Building: ML classification methods will be utilized in this effort to identify the crops. XGBoost is driven by a dynamic artificial fish swarm used as a model-building method.

v. Model Evaluation: Accuracy, recall, precision, f1-score, training time, and MSE parameters are employed to evaluate the simulation outcomes. Figure 1 illustrates the procedural flow of the DAFS-XGBoost model.

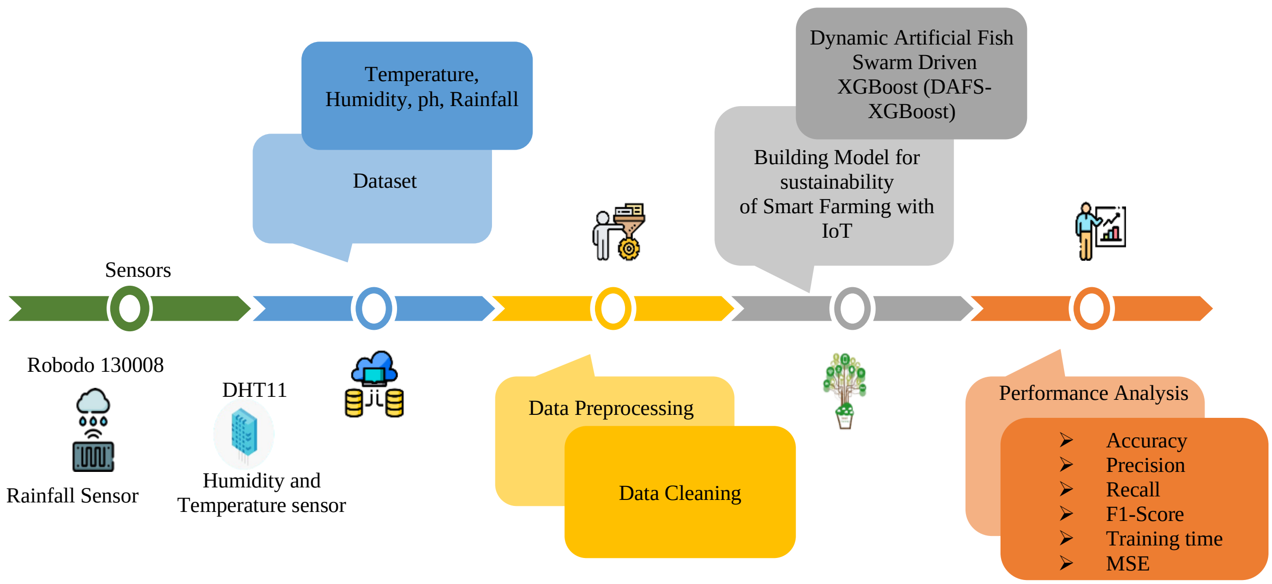

Figure 1: Proposed DAFS-XGBoost model

Data Collection

The datasets were obtained from the Kaggle website. The dataset utilized by the crop recommendation component is called recommender.csv. The datasets utilized in this module are summarized in Table 2 and Figure 2. The crop recommendation was utilized to train the model because it includes characteristics that are necessary for humidity, potassium requirement ratio, crop prediction, soil pH, phosphorous requirement ratio, nitrogen need ratio, temperature, and average rainfall.

Table 2: Details of the crop suggestion module's dataset

Attributes	Crop_Recommendation.csv
Source	https://www.kaggle.com/code/joanpau/iot-farming-crops-soils/input
No. of samples	8
Attributes	Classification
Used for	PyTorch
Labels count	22

Index	N	P	K	temperature	humidity	ph	rainfall	label
0	90	42	43	20.879744	82.002744	6.502985	202.935536	rice
1	85	58	41	21.770462	80.319644	7.038096	226.655537	rice
2	60	55	44	23.004459	82.320763	7.840207	263.964248	rice
3	74	35	40	26.491096	80.158363	6.980401	242.864034	rice
4	78	42	42	20.130175	81.604873	7.628473	262.71734	rice
...	...	...	...	...	...	...	...	...
2195	107	34	32	26.774637	66.413269	6.780064	177.774507	coffee
2196	99	15	27	27.417112	56.636362	6.086922	127.92461	coffee
2197	118	33	30	24.131797	67.225123	6.362608	173.322839	coffee
2198	117	32	34	26.272418	52.127394	6.758793	127.175293	coffee
2199	104	18	30	23.603016	60.396475	6.779833	140.937041	coffee

Figure 2: Dataset sample

Data Cleaning

The process of correcting inaccuracies, correcting errors and inconsistencies, in a dataset such that it is reliable and of high quality for modeling or evaluation is referred to as information cleansing. Enhance the usefulness of the dataset, this entails performing operations that include resolving missing values, eliminating duplicates, and correcting errors. In this context, data cleaning methods are employed for preprocessing, ensuring that the dataset is prepared effectively for subsequent analysis and model training.

Descriptive analyses not only provide a framework for the dataset's appearance but also assist in the discovery of new information. Initially, the dataset was imported and the missing values for each attribute were analyzed. The characteristics in the dataset are missing values-free; the results are displayed in Figure 3 (a). If it has been established that there are no missing values, Figure 3 (b) shows the data format of the properties, and Figure 3 (c) depicts the unique values in the dependent factor, which is the labeling component.

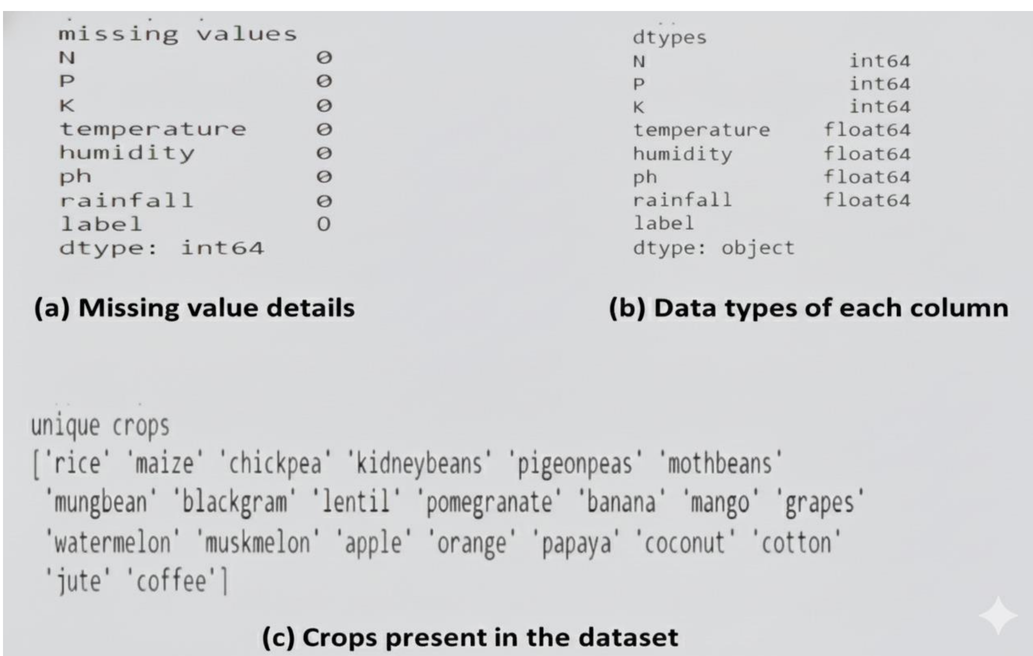

Figure 3: Dataset sample values

Splitting the Data into Train and Test Sets

There are testing and training sets within the dataset as the first phase of method development. Data for training is the first set of information needed to train ML techniques. Training datasets are used to teach machine learning algorithms that generate precise predictions or complete a specified task. A set of tests is a collection of information that is intended to give a precise assessment of the final model fit on the training sample. Research split data into 70:30 testing and training sets.

Model Building for Dynamic Artificial Fish Swarm Driven XGBoost

This effort leverages machine learning classification techniques to advance smart farming practices, aiming to enhance both productivity and sustainability. The combination of IoT technology with the Dynamic Artificial Fish Swarm Driven XGBoost (DAFS-XGBoost) model plays a pivotal role in achieving this goal. By harnessing the robust predictive capabilities of XGBoost, combined with the dynamic optimization provided by artificial fish swarm optimization, the DAFS-XGBoost method introduces significant improvements to traditional farming techniques. This advanced approach allows for real-time adaptation to varying farming conditions, optimizing resource use while maximizing crop yields and minimizing environmental impacts.

The DAFS-XGBoost model performs the important role of integration with IoT sensors like temperature, humidity of the soil, and so on to produce insights for farmers. Such integration makes it possible to have precision agriculture due to the accurate decisions that are made. The model updates the crop management and resource and decision-making processes in real-time, and it provides for early interventions to such threats as pests, diseases, and adverse weather conditions. This way farmers can use more accurate farming approaches and manage problems before they advance, thus increasing the sustainability and effectiveness of farm practices.

By combining XGBoost with Dynamic Artificial Fish Swarm (DAFS) optimization, the suggested approach improves smart farming. This dynamic adjustment is more balanced compared to the examination and exploitation during hyperparameter tuning and optimizes depth to maximum and learning rate hyperparameters. XGBoost comes out very handy in smart farming applications due to its ability to handle big data and high speeds. The technique is applied for IoT sensor data and addresses specific challenges in the agriculture domain. Due to its flexibility and scalability to the volume of data and operational conditions, it might be applied in a wide variety of agriculture-related processes.

Dynamic Artificial Fish Swarm Algorithm (DAFS)

The DAFS optimizes utility allocation in smart farming by simulating fish behavior to distribute responsibilities effectively. It adapts to changing environmental conditions, enhancing crop management, irrigation scheduling, and pest manipulation. Two main modifications were made to the DFAS in this study. It has an increased ability to identify the best fish swarm and dynamic vision.

Dynamic Vision The vision factor is essential to the DAFS process because it predicts the amounts of surrounding fish, the intended fish are communicating with, which has a significant impact on the efficiency of the swarm and subsequent processes. Increasing the eyesight parameter improves the chances of discovering fish with greater FV; however, it may also force the fish swarm to congregate in one place. Furthermore, as species can quickly settle into a local optimum, this tends to reduce species variety.

Lowering the eyesight parameter induces the social group to disperse and lessens the amount of surrounding fish, but it also increases species diversity. This expands the search space and raises the possibility of discovering the best solution, but it also lengthens the time needed to reach confluence. It might be difficult to achieve an appropriate equilibrium when it comes to assigning the eyesight parameter. This study created a system known as dynamic vision to get around this problem. It assigns each fish a set of visual parameters based on its unique circumstances. For example, a fish with a lower FV needs to see more clearly to locate a solution quickly. In the different combinations, fish having a greater level of fitness can improve local searches by using a smaller vision parameter.

Each fish determines the value of its visual metric depending on its FV. For example, if the FV is higher than average, the eyesight metric is dropped, and vice versa. This is how the endocrine-based equation is calculated:

$$EM(j) = e_1\left(\frac{e_{max}-e_j}{e_{max}-e_{avg}}\right).[\frac{\pi}{2} + e_2\left(e_j\,\frac{e_{j-1}-e_{j+1}}{2}\right) \quad (1)$$

$$vision\ (j) = vision\ (static).EM\ (j).CV, \quad (2)$$

Here the fish endocrine system is represented by $EM(j)$. The variables e_j, e_{max}, and e_{avg}denote the maximum and average FV of the fish inside the school, respectively.

The FV of fish $e_{j-1}$are represented bye_{j-1}, whereas those of fish $e_{j+1}$are represented by e_{j+1}. To correct the endocrine system's range, $e_1(w) = \mathrm{atan}(w)$, $e_2(w) = \mathrm{atan}(-w)$. The typical separation among every fish is used to derive the initial measurement in Equation (2) and CV is the correction constant. Regardless of their FV, fish j and i would have the same eyesight under the original vision variable. Fish $j's$ FV is increased to 80 after that is adjusted using the endocrine-based formula, which impairs its vision but improves its capacity to conduct searches locally. Fish $i's$ FV, in contrast, decreased to 65, improving its eyesight and ability to conduct global searches to find solutions more quickly.

Searching for the Best Fish Swarm To increase the capability of nearby searches and prevent settling into an optimal location, this study used a straightforward technique of looking for the best fish swarm. To increase the probability of discovering the ideal solution, the procedure replicates the fish through the highest FV into a fish swarm when each iteration is complete. This modification was made possible by the addition of four parameters: MCN, BMR, BSN, and BFN.

BFN is the quantity of fish that were duplicated. The 5 best fish, for instance, would be copied to the BFN = 5. The LS number for every fish in the best fish swarm is represented. The algorithm would carry out local ten searches for every fish in the best fish swarm if BSN = 10. The BMR represents the rate of evolution of LS, while the MCN of feature modifications during every regional search is denoted by the MCN. Consider the following scenario: let's say that the feature number was 8, the BMR was 0.1, and the MCN was set to 1, 8 $*$ 0.1 = 0.8. In this instance, 0.8 is less than MCN; therefore, in each LS for the best fish swarm, one random alteration would be made. After five LS, the fish with the greatest feature subgroup from this group of highly ranking fish replaces the fish that was previously declared the best fish.

Extreme Gradient Boosting Algorithm (XGBoost)

XGBoost is utilized in SF to predict crop yields based on various environmental factors like soil moisture, temperature, and rainfall. XGBoost can optimize crop management decisions, which include irrigation scheduling and fertilizer software, main to stepped forward yields and resource efficiency. The GBDT structure serves as the foundation for the XGBoost algorithm. It deserves a great deal of attention due to its exceptional performance in Kaggle's machine-learning competitions. The regularized term in the XGBoost goal function prevents over-fitting, in contrast to GBDT. Here's an explanation of the major goal operate (see Equation (3)).

$$P = \sum_{j=1}^{m} K(\, z_j, E(w_j)) + \sum_{l=1}^{s} Q(\, e_l) + D \tag{3}$$

Here $Q(e_l)$is the term that regulates at repetition l and D is a constant that can be deleted precisely. The phrase for regularization $Q(e_l)$is expressed as Equation (4).

$$Q(e_l) = \alpha G + \frac{1}{2}\eta \sum_{l=1}^{G} \omega_i^2 \tag{4}$$

Here G is the number of leaves, η is the consequence factor, and ω_i is the productivity outcome for each leaf node. $\alpha$is the complexity of leaves. According to the classification criteria, leaves indicate the expected categories, whereas the leaf node represents the undividable tree node.

Furthermore, instead of using the initial-order derivative, XGBoost uses the Taylor series of second-order major functions, in contrast to GBDT. The function can be communicated as follows if the MSE serves as the loss value:

$$O = \sum_{j=1}^{m} \left[oj^{\omega} r^{(w_j)} + \frac{1}{2}\left(rj^{\omega^2} r(w_j) \right) \right] + \alpha G + \frac{1}{2}\eta \sum_{i=1}^{G} \omega_i^2 \tag{5}$$

Where $h_j$and $g_j$stand for the initial and secondary elements of the degradation function and $r(w_j)$is a feature that converts points of information into leaf mappings.

The total amount of loss is the total of each of the loss rates. Because observations in the DT relate to leaf nodes, the final loss value might be calculated by adding the total losses of the leaf nodes. Consequently, the following is an expression of the fundamental function:

$$O = \sum_{i=1}^{S} \left[oj^{\omega} r + \frac{1}{2}(R_i + \eta)\omega_i^2) \right] + \alpha G \tag{6}$$

Where $O_i = \sum_{j \epsilon J_i} o_j, R_i = \sum_{j \epsilon J_i} r_j \, , and \; J_i$represent. The overall amounts of specimens in each leaf node I. To sum up, the process of maximizing the main purpose is made easier by locating the minimum value of an exponential function. XGBoost can reduce over-fitting more successfully if normalization processes are incorporated.

3 Results and Discussion

Experimental Setup

The experimental setup for validating the Dynamic Artificial Fish Swarm Driven XGBoost (DAFS-XGBoost) model involves several key steps. IoT sensors, including the DHT11 for temperature and humidity and the Robodo 130008 Raindrops Detection Sensor, collect environmental data such as temperature, humidity, and rainfall. This data is preprocessed to ensure accuracy before being used to train the XGBoost model with MATLAB's Statistics and Machine Learning Toolbox. To support the computational needs, a minimum of 8 GB of RAM is recommended. The hyperparameters and their corresponding superior values of the proposed approach are offered in Table 3. Table 4 shows the experimental setup.

The metrics of precision, accuracy, recall, and MSE are examined in this section. The effectiveness of the classification of the NB, KNN, RF (Sharma et al., 2023), RFR, SVR, MLR (Pawar et al., 2021) and KNN, DT, and NB (Mathi et al., 2023) classifiers is being compared. This comparison provides insights into the effectiveness and efficiency of the suggested method relative to other common ML techniques. The setup ensures that the DAFS-XGBoost model is thoroughly evaluated for both predictive accuracy and computational effectiveness in various agricultural contexts. Figure 4 depicts the graph illustrating the accuracy loss of the suggested approach.

Table 3: Parameters and values of proposed methods

Parameters	Values
MCN	1
BSN	20
Maximum number of prey	20
G	[0.00001,8]
BMR	0.15
BFN	5
CV	0.5
Fish number	30
C	[0.01,1024]
Maximum crowd degree	0.5
n estimators	133
Learning rate	0.03
Maximum depth	4

Table 4: Hardware and software experimental setup

Component	Details
Hardware	Computer: Minimum 8 GB of RAM
Software	Sensors: DHT11 (temperature and humidity), Robodo 130008 Raindrops Detection Sensor MATLAB software

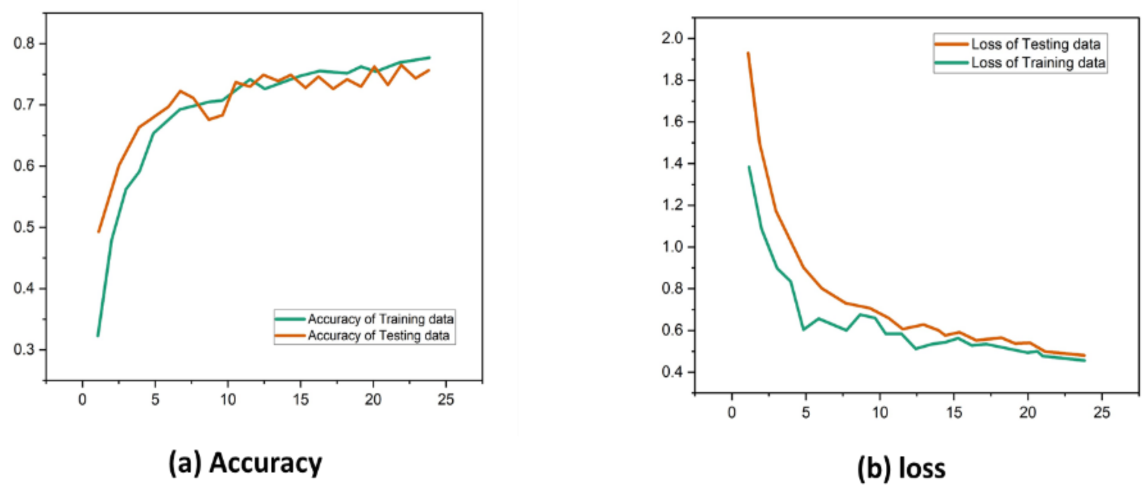

Figure 4: Accuracy and loss

The percentage of all farming circumstances (such as temperature, soil moisture, and crop health) that the network driven on the IoT accurately detects out of all cases is known as accuracy. It is crucial for guaranteeing the best possible farming techniques and control of resources, and it represents the entire system's efficacy and dependability in producing accurate forecasts.

Figure 5 illustrates the accuracy rate achieved by the proposed methodology. Compared to other methods, the suggested model achieves an accuracy rate of 96.73%. In comparison, naive Bayes, KNN, and RF had accuracy rates of 94.23%, 95.85%, and 93.45%, respectively. DAFA-XGBoost has superior outcomes compared to other existing methods.

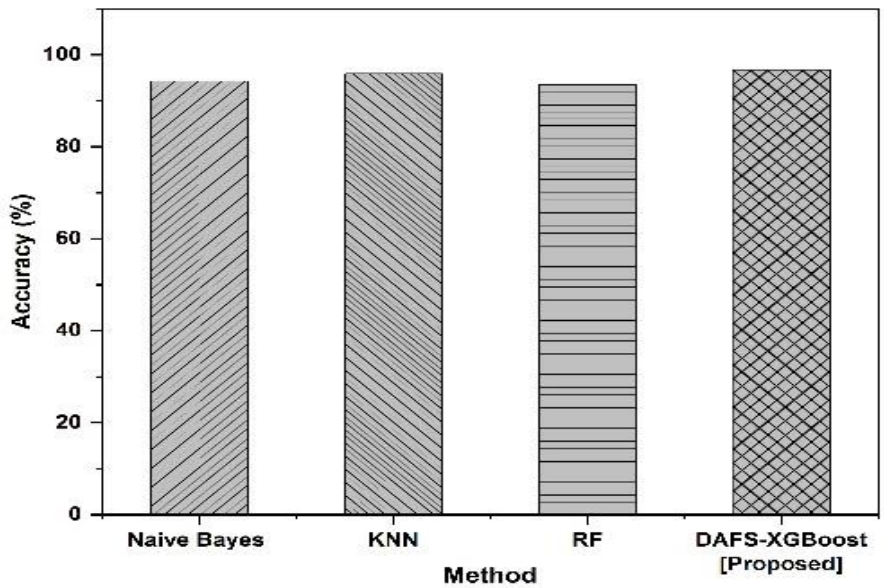

Figure 5: Results of Accuracy

Out of all conditions projected to be positive, precision indicates the percentage of actual positive farming conditions found by the IoT sensors. Precision assesses the number of the method's predictions, for instance, that a crop needs to be watered, that are accurate. High precision is essential to smart farming to reduce needless operations that could waste resources, such as over-watering.

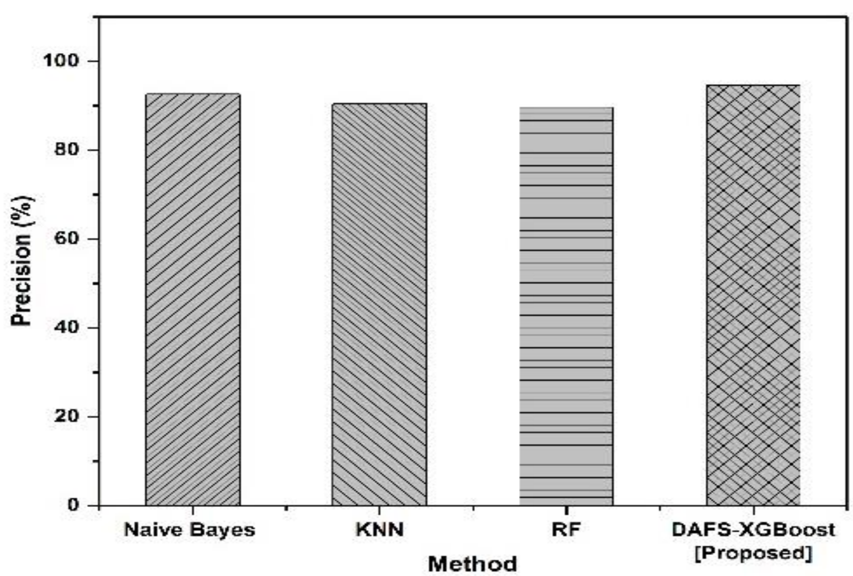

Figure 6: Results of Precision

The precision rate that the suggested methodology was able to obtain is shown in Figure 6. The suggested model achieves a precision rate of 94.61% in comparison to other methods. Naive Bayes (NB), KNN, and RF, in contrast, had precision rates of 92.59%, 90.45%, and 89.67% respectively. When compared to other existing methods, DAFA-XGBoost produced better results.

Recall measures the percentage of real positive farming circumstances that the IoT network accurately detects out of all real positive situations. It shows how well the system can identify important farming requirements like outbreaks of diseases or infestations of pests. A high recall guarantees that the IoT network can detect the majority of positive cases, which is essential for prompt farming solutions.

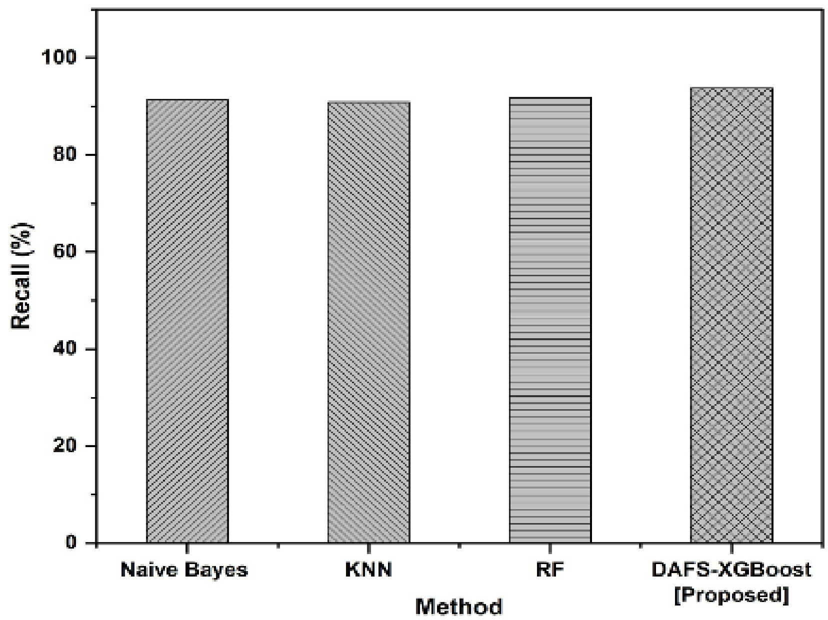

Figure 7: Results of recall

The recall that the suggested methodology was able to acquire is shown in Figure 7. The proposed model achieves a 93.76% recall rate in comparison to other methods. By comparison, the recall rates of naive Bayes, KNN, and RF were 91.32%, 90.81%, and 91.76% respectively. When DAFA-XGBoost was compared to other current approaches, better results were obtained. Table 5 depicts the values of recall, accuracy, and precision.

Table 5: Values of recall, accuracy, and precision

Method	Accuracy (%)	Precision (%)	Recall (%)
NB	94.23	92.59	91.32
KNN	95.85	90.45	90.81
RF	93.45	89.67	91.76
DAFS-XG Boost [proposed]	96.73	94.61	93.76

A metric called MSE evaluates the mean of the squared deviations between the observed and anticipated values. When it comes to smart farming, MSE assesses how accurate the IoT system’s ongoing forecasts are, such as temperature or soil moisture content. For accurate farming operations and decision-making, a lower MSE means that the method's forecasts are closer to the actual values.

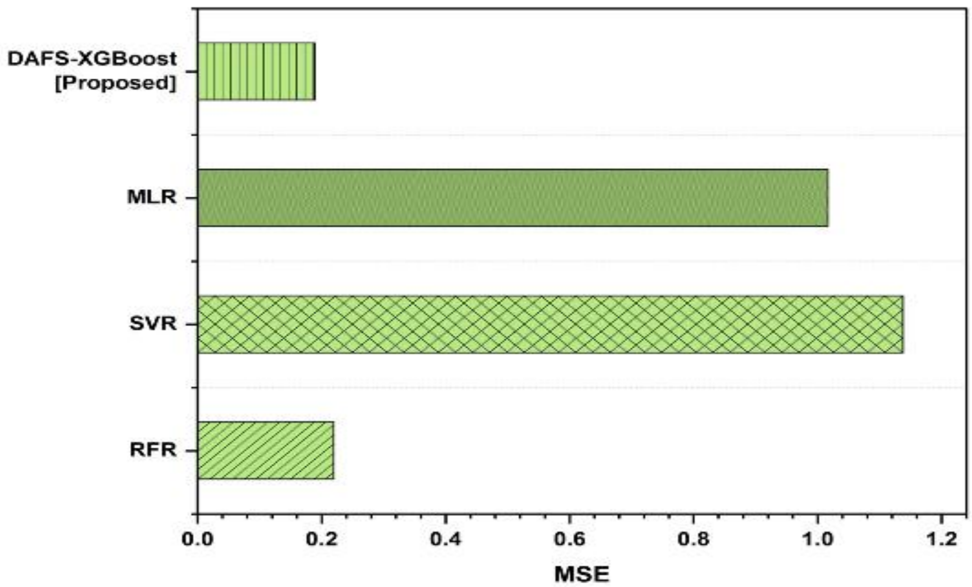

Figure 8: Results of MSE

Figure 8 displays the MSE value that the recommended methodology was able to obtain. The suggested method achieved 0.19 of MSE. In contrast, RFR, SVR, and MLR all had MSE values of 0.22, 1.138, and 1.0168. Better results were observed when DAFA-XGBoost was compared to other existing techniques. Table 6 demonstrates the values of MSE.

Table 6: Values of MSE

Method	MSE
RFR	0.22
SVR	1.138
MLR	1.0168
DAFS-XG Boost [proposed]	0.19

By calculating the harmonic average of both precision and recall, the F1-score offers a statistic that equalizes the trade-off between the two. This statistic is very helpful in smart farming when there are serious effects from both false positives and false negatives, like in crop disease prediction. One method for evaluating the total accuracy is the F1-score of the IoT system in controlling farming activities.

Figure 9 displays the F1-score value that the recommended methodology was able to obtain. The suggested method achieved 90.12% of the F1 score. In contrast, KNN, DT, and NB all had F1-score values of 67.11%, 57.32%, and 64.46%. Better results were observed when DAFA-XGBoost was compared to other existing techniques.

The time needed to train machine learning algorithms that drive the IoT-driven smart farming system is referred to as training time. Because it influences how quickly novel systems can be updated or deployed, this measure is significant. Reducing training time is crucial to keeping the IoT network current and responsive in an evolving farming environment where conditions might change quickly.

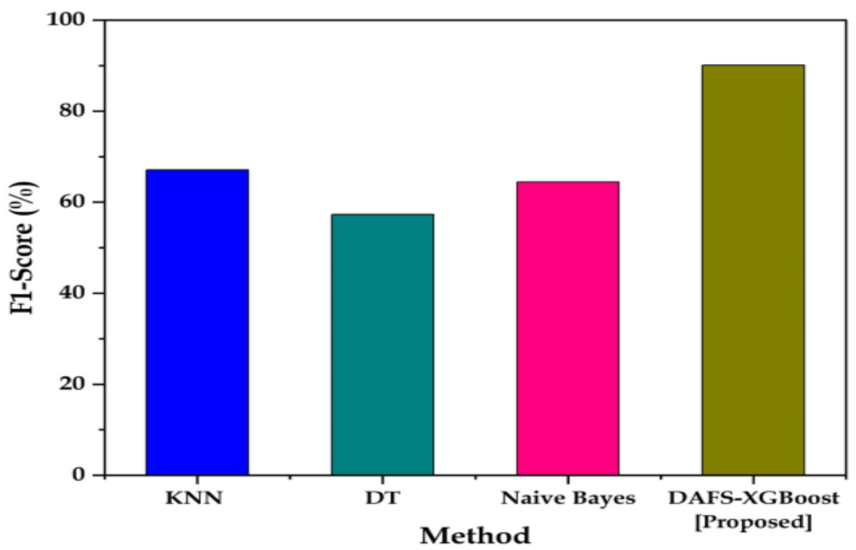

Figure 9: Results of F1-score

Figure 10 displays the training time value in seconds that the recommended methodology was able to obtain. The suggested method achieved 0.0006 of training time. In contrast, KNN, DT, and NB all had training time values of 0.0008, 0.0032, and 0.0033. Better results were observed when DAFA-XGBoost was compared to other existing techniques. Table 7 shows the values of the F1-score and training time.

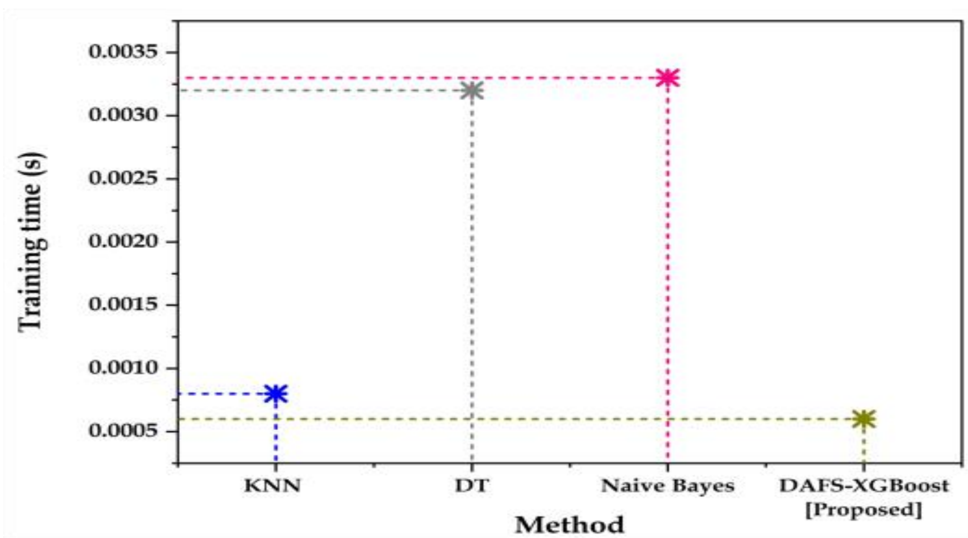

Figure 10: Results of training time

Table 7: Values of F1-score and training time

Method	F1-score (%)	Training time (s)
NB	67.11	0.0008
KNN	57.32	0.0032
RF	64.46	0.0033
DAFS-XG Boost [proposed]	90.12	0.0006

Large datasets may cause RF's computational complexity to be constrained in smart farming, resulting in slower predictions. Its performance could be impacted by its inability to handle extremely unbalanced data and comprehend features. For example, RF obtained 93.45% accuracy, 89.67% precision, and 91.76% recall. These results are strong, but they might not always meet the efficiency and precision needed for real-time applications. Naïve Bayes assumes feature independence, which oversimplifies complicated connections among agricultural data; it may not perform well in smart farming scenarios. Its 94.23% accuracy, 92.59% precision, 91.32% recall, 64.46% F1-score, and 0.0033 second training time all clearly demonstrate this restriction. Prediction accuracy and reliability may suffer as a result of such performance, particularly when dealing with multi-dimensional datasets and linked environmental factors.

KNN's performance metrics are accuracy of 95.85%, precision of 90.45%, recall of 90.81%, F1-score of 67.11%, and training time of 0.0008 seconds, which makes clear the network's shortcomings in smart farming. These numbers demonstrate how KNN has a high computational cost and is susceptible to noise that can compromise accuracy and decision-making in big or complicated datasets. Overfitting of DT in smart farming can result in subpar generalization. They are also less effective at handling large-scale or high-dimensional data and have trouble grasping complex relationships. With an F1-score of 57.32% and a training time of 0.0032 seconds in the study's setting, DT's limitations in terms of efficiency and performance when compared to more advanced techniques are made clear.

In smart farming, RFR may have limitations due to its high computational expense and subpar real-time performance. It might also have trouble correctly capturing complex feature interactions in dynamic agricultural contexts, given its MSE value of 0.22. The application of SVR in smart farming may be constrained by its high computing complexity and tuning sensitivity. With an MSE of 1.138, SVR can find it difficult to manage complicated relationships and big datasets. Smart farming may be limited by the simplicity of MLR, which might result in the underfitting of intricate data patterns. Because of its MSE value of 1.0168, MLR may not be able to accurately predict complex feature interactions in dynamic situations.

DAFS-XGBoost addresses these limitations effectively by offering superior accuracy (96.73%), precision (94.61%), recall (93.76%), and F1-score (90.12%), while maintaining a lower MSE (0.19) and faster training time (0.0006 seconds). In comparison to other approaches, this makes it a strong choice for maximizing smart farming, offering more precise, effective, and responsive performance.

Implications

The above outcomes show how the DAFS-XGBoost approach gains over a number of the shortcomings of the present models. The fact that the software is beneficial for complex agricultural information is proved by its high accuracy, precision rate, and recall index, as well as low MSE. Furthermore, owing to its shorter training time, model updates are also ensured at a very high rate. For this reason, the DAFS-XGBoost is well-suited to real-time smart farming scenarios where real-time decision-making is critical. DAFS-XGBoost therefore shows a better approach to the problem of efficiency in smart farming, overcoming the limitations of such techniques as reduced recall and precision, increased mean square error, and extended training times. It may lead to higher production of crops, efficient and productive uses of resources, and more efficient farming techniques; all factors that could lead to the growth and sustainability of the farming sector.

4 Conclusion

This study introduces a novel technique for smart farming leveraging IoT sensor data incorporated with the modern Dynamic Artificial Fish Swarm Driven XGBoost (DAFS-XGBoost) model. Utilizing a Kaggle dataset, specifically the "Crop recommender.csv" file, which encompasses critical IoT sensor metrics including temperature, humidity, rainfall, soil pH, and nutrient ratios, we implemented data cleaning processes by dividing the dataset into subsets for testing and training. The proposed DAFS-XGBoost methodology was implemented via MATLAB. Our findings emphasize the efficacy of the DAFS-XGBoost technique, as demonstrated by its advanced overall performance with recall at 93.76%, accuracy at 96.73%, precision at 94.61%, F1-score 90.12%, training time of 0.0006 and MSE at 0.19. Using an IoT-assisted framework in conjunction with a sustainable smart farming method greatly increases agricultural output and the efficiency of resources. Real-time monitoring and data collection are made possible by the integration of IoT sensors, which promotes improved farming methods and well-informed decision-making. By anticipating and managing environmental variables more skillfully, the DAFS-XGBoost model enhances prediction accuracy even further, supporting sustainability and cutting waste. The method is not without its drawbacks, though. Model performance may be impacted by potential sensitivity to parameter adjustment, and scaling issues with big datasets may prevent the model from being used in more general scenarios. Additionally, model tuning may become questionable if heuristic optimization is relied upon. To overcome these obstacles, future research may combine many meta-heuristic algorithms for increased efficiency, investigate uses other than classification, and create plans to reduce overfitting in intricate datasets.

References

[1] Alrowais, F., Asiri, M. M., Alabdan, R., Marzouk, R., Hilal, A. M., & Gupta, D. (2022). Hybrid leader-based optimization with deep learning-driven weed detection on the Internet of Things enabled smart agriculture environment. *Computers & Electrical Engineering, 104*, 108411. https://doi.org/10.1016/j.compeleceng.2022.108411

[2] Anand, T., Sinha, S., Mandal, M., Chamola, V., & Yu, F. R. (2021). AgriSegNet: Deep aerial semantic segmentation framework for IoT-assisted precision agriculture. *IEEE Sensors Journal, 21*(16), 17581–17590. https://doi.org/10.1109/JSEN.2021.3071290

[3] Bahmaid, S., & Ghaleb, S. A. M. (2024). Intrusion Detection System Using Chaotic Walrus Optimization-based Convolutional Echo State Networks for IoT-assisted Wireless Sensor Networks. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 15*(3), 236-252. https://doi.org/10.58346/JOWUA.2024.I3.016

[4] Charania, & Li, X. (2020). Smart farming: Agriculture's shift from a labor-intensive to technology-native industry. *Internet of Things, 9*, 100142. https://doi.org/10.1016/j.iot.2019.100142

[5] Dahane, R., Benameur, R., & Kechar, B. (2022). An IoT low-cost smart farming for enhancing irrigation efficiency of smallholder farmers. *Wireless Personal Communications, 127*(4), 3173–3210. https://doi.org/10.1007/s11277-022-09915-4

[6] Dhanaraju, M., Chenniappan, P., Ramalingam, K., Pazhanivelan, S., & Kaliaperumal, R. (2022). Smart farming: Internet of Things (IoT)-based sustainable agriculture. *Agriculture, 12*(10), 1745. https://doi.org/10.3390/agriculture12101745

[7] Gupta, & Nahar, P. (2023). Sandpiper optimization algorithm with cosine similarity based cross-layer routing protocol for smart agriculture in wireless sensor network assisted Internet of Things systems. *International Journal of Communication Systems, 36*(11), e5514. https://doi.org/10.1002/dac.5514

[8] Idoje, G., Dagiuklas, T., & Iqbal, M. (2021). Survey for smart farming technologies: Challenges and issues. *Computers & Electrical Engineering, 92*, 107104. https://doi.org/10.1016/j.compeleceng.2021.107104

[9] Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2022). Enhancing smart farming through the applications of Agriculture 4.0 technologies. *International Journal of Intelligent Networks, 3*, 150–164. https://doi.org/10.1016/j.ijin.2022.09.004

[10] Khan, A. I., Alsolami, F., Alqurashi, F., Abushark, Y. B., & Sarker, I. H. (2022). Novel energy management scheme in IoT-enabled smart irrigation system using optimized intelligence methods. *Engineering Applications of Artificial Intelligence, 114*, 104996. https://doi.org/10.1016/j.engappai.2022.104996

[11] Kumar, T. M. S. (2024). Security challenges and solutions in RF-based IoT networks: A comprehensive review. *SCCTS Journal of Embedded Systems Design and Applications, 1*(1), 19-24.

[12] Lytos, A., Lagkas, T., Sarigiannidis, P., Zervakis, M., & Livanos, G. (2020). Towards smart farming: Systems, frameworks, and exploitation of multiple sources. *Computer Networks, 172*, 107147. https://doi.org/10.1016/j.comnet.2020.107147

[13] Mahajan, H. B., Badarla, A., & Junnarkar, A. A. (2021). CL-IoT: Cross-layer Internet of Things protocol for intelligent manufacturing of smart farming. *Journal of Ambient Intelligence and Humanized Computing, 12*(7), 7777–7791. https://doi.org/10.1007/s12652-020-02502-0

[14] Mahbub, M. (2020). A smart farming concept based on smart embedded electronics, the Internet of Things, and a wireless sensor network. *Internet of Things, 9*, 100161. https://doi.org/10.1016/j.iot.2020.100161

[15] Mathi, S., Akshaya, R., & Sreejith, K. (2023). An Internet of Things-based efficient solution for smart farming. *Procedia Computer Science, 218*, 2806–2819. https://doi.org/10.1016/j.procs.2023.01.252

[16] Mohamed, E. S., Belal, A. A., Abd-Elmabod, S. K., El-Shirbeny, M. A., Gad, A., & Zahran, M. B. (2021). Smart farming for improving agricultural management. *The Egyptian Journal of Remote Sensing and Space Science, 24*(3), 971–981. https://doi.org/10.1016/j.ejrs.2021.08.007

[17] Monica Nandini, G.K. (2024). Iot Use in A Farming Area to Manage Water Conveyance. *Archives for Technical Sciences, 2*(31), 16-24. https://doi.org/10.70102/afts.2024.1631.016

[18] Nambiar, K., & Varma, A. (2023). A Comprehensive Analysis of Smart Farming Using Wireless Sensor Networks. *International Academic Journal of Science and Engineering, 10*(3), 12–16. https://doi.org/10.71086/IAJSE/V10I3/IAJSE1025

[19] Neethu, M. C., & Ramyaprabha, N. (2025). A Study on Tourist Push and Pull Motives to Visit Farm Tourism Sites in Kerala. *Indian Journal of Information Sources and Services, 15*(1), 1–5. https://doi.org/10.51983/ijiss-2025.IJISS.15.1.01

[20] Novak, P., & Vacek, T. (2023). An IoT-Driven Predictive Maintenance Model for Smart Manufacturing Systems. *International Academic Journal of Innovative Research, 10*(3), 9–15. https://doi.org/10.71086/IAJIR/V10I3/IAJIR1019

[21] Pal, S., Vijay Kumar, H., Akila, D., Jhanjhi, N. Z., Darwish, O. A., & Amsaad, F. (2023). Information-centric IoT-based smart farming with dynamic data optimization. *Computers, Materials & Continua, 74*(2), 3865–3880. http://dx.doi.org/10.32604/cmc.2023.029038

[22] Patil, N., & Khairnar, V. D. (2021). IoT based smart farming system. *International Journal of Advanced Research Ideas and Innovations in Technology, 7*(1).

[23] Pawar, S., Dere, S., Akangire, A., Kamble, H., & Shrawne, S. (2021). Smart farming using machine learning. *Smart Computing*, 73–81.

[24] Sharma, Sharma, Tselykh, A., Bozhenyuk, A., Choudhury, T., Alomar, M. A., & Sánchez-Chero, M. (2023). Artificial intelligence and Internet of Things oriented sustainable precision farming: Towards modern agriculture. *Open Life Sciences, 18*(1), 20220713. https://doi.org/10.1515/biol-2022-0713

[25] Singh, G., & Singh, J. (2023). A cost effective IoT-assisted framework coupled with fog computing for smart agriculture. In *2023 IEEE 8th International Conference for Convergence in Technology (I2CT)* (1–8). https://doi.org/10.1109/I2CT57861.2023.10126231

[26] Sukanya, V., & Ramachandram, S. (2023). A novel all members group search optimization based data acquisition in cloud assisted wireless sensor network for smart farming. *International Journal of Computer Networks and Applications*, 651–667. https://doi.org/10.22247/ijcna/2023/223318

Authors Biography

Ashvin B Amale is currently serving as an Assistant Professor in the Department of Mechanical Engineering at Yeshwantrao Chavan College of Engineering, Nagpur. His areas of expertise include thermal engineering, renewable energy systems, and computational fluid dynamics. He is actively involved in academic research and project supervision, with a strong focus on industry-academia collaboration and engineering education development.

Dr.K.V. Pradeep is an Associate Professor in the School of Computer Science and Engineering at Vellore Institute of Technology (VIT), Chennai campus. His teaching and research interests span across artificial intelligence, machine learning, data analytics, and software engineering. He has published several research papers in reputed journals and conferences and is actively involved in mentoring students and leading funded research initiatives.

Rishabh Bhardwaj is associated with the Centre of Research Impact and Outcome, Chitkara University, Rajpura, Punjab, India. His research focuses on research analytics, institutional performance measurement, and strategies to enhance academic impact and visibility. He is engaged in projects that emphasize data-driven approaches to research assessment and evidence-based improvement of institutional research practices. His contributions aim to strengthen innovation, collaboration, and research excellence across academic environments.

Dr. Neeta Mazumdar is an Associate Professor in the Department of Mathematics at the Government College of Arts, Science and Commerce, Quepem, Goa. With significant academic and research experience, her expertise spans pure and applied mathematics. She is actively involved in mathematical modeling, numerical analysis, and interdisciplinary research. Dr. Mazumdar contributes to curriculum development and promotes research-oriented learning among students. Her scholarly work is reflected in various national and international publications.

Sachin S Pund is an Assistant Professor in the Department of Mechanical Engineering at Shri Ramdeobaba College of Engineering and Management, Nagpur. His academic and research interests include thermal engineering, renewable energy systems, and advanced manufacturing technologies. He has contributed to various academic publications and actively participates in curriculum development, research supervision, and professional training programs.

Bhuvana Jayabalan is an Associate Professor in the Department of Computer Science and Information Technology at JAIN (Deemed-to-be University), Bangalore. With a strong academic foundation and research orientation, her areas of expertise include data analytics, artificial intelligence, and machine learning. She has authored and co-authored several research papers in reputed journals and actively mentors students in academic and industry-oriented projects. Dr. Bhuvana is committed to promoting research excellence and innovative practices in higher education.