



Modified Salp Swarm Algorithm with Adaptive Weighting Based Bidirectional LSTM Network Ensemble Method for Crop Recommendation

Parijata Majumdar^{1,*}

¹ Indian Institute of Information Technology Agartala, Barjala, Tripura 799046, India

Abstract

Farmers sometimes grow crops with low yields, wasting land, labor, and time—especially in developing countries where demand for food is increasing. A Crop Recommendation System (CRS) can help by using precision farming techniques that analyze soil and environmental data to suggest the most suitable crops. This study proposes a CRS using a Modified Salp Swarm Algorithm (MSSA) for feature selection and an Adaptive Weighted Bi-directional Long Short-Term Memory (AWBiLSTM) ensemble for prediction. MSSA enhances the original algorithm by improving local search and convergence speed, addressing SSA's limitations. Climate data is pre-processed and relevant features are selected using MSSA. AWBiLSTM then predicts suitable crops with improved accuracy. Experimental results show that the MSSA-AWBiLSTM approach outperforms existing methods in precision, recall, and execution time. The proposed method obtains an accuracy of 98.72%, precision of 98.81%, recall of 98.54%, specificity of 98.10%, PR-Score of 98.18%, ROC-Score

of 98.49%, F1-Score of 98.48%, and Matthews Correlation Coefficient (MCC) of 97.63%.

Keywords: soil and climate conditions, feature Selection, crop recommendation, modified salp swarm algorithm, adaptive weighting based BiLSTM.

1 Introduction

Precision Agriculture (PA) utilizes advanced sensors and analytical tools to improve crop yield and support decision-making [1]. Globally adopted, PA helps reduce labor, optimize resource use, and improve the management of fertilizers and irrigation. Successful crop selection depends on several factors including irrigation, rainfall, soil conditions, sunlight, and pest control [2, 20]. One major challenge is choosing crops compatible with the soil conditions [3]. Climate and soil factors must be thoroughly analyzed before implementing any PA strategy, as they directly impact crop growth [4, 5].

Accurate crop recommendations are essential to avoid losses many farmers cannot afford [6]. However, traditional machine learning (ML) models often produce inaccurate results due to noisy and incomplete datasets [7, 8]. Deep learning (DL) models have emerged as a more effective alternative, capable of training on large datasets and solving complex decision



Submitted: 23 April 2025

Accepted: 04 May 2025

Published: 29 May 2025

Vol. 1, No. 1, 2025.

10.62762/TSEL.2025.947593

*Corresponding author:

✉ Parijata Majumdar

parijata.majumdar@iiitagartala.ac.in

Citation

Majumdar, P. (2025). Modified Salp Swarm Algorithm with Adaptive Weighting Based Bidirectional LSTM Network Ensemble Method for Crop Recommendation. *ICCK Transactions on Swarm and Evolutionary Learning*, 1(1), 3–11.

© 2025 ICCK (Institute of Central Computation and Knowledge)

problems with greater accuracy [9]. The Internet of Things (IoT) further strengthens PA by enabling real-time data collection and analysis. Climate information is critical for planning, and timely crop selection is key to maximizing yield [10]. Classification models using variables like temperature, humidity, soil type, pH, and crop period can support accurate decision-making [12].

Farmers still often rely on intuition due to limited access to real-time, reliable data, which leads to poor crop choices and reduced productivity [11]. Despite existing research, crop recommendation systems (CRS) still face issues with data quality, long processing times, and low accuracy. To overcome these limitations, DL models have been applied to improve both efficiency and precision [11]. Among optimization methods, the Salp Swarm Algorithm (SSA), inspired by marine salps, has shown success in various domains like feature selection and forecasting [13–15, 17]. SSA benefits from minimal parameter tuning and strong neighborhood search but can face slow convergence and local optima stagnation [16].

To improve SSA's performance, we introduce a Modified SSA (MSSA) that incorporates local best solution strategies, enhancing population diversity and convergence. This is paired with an Adaptive Weighted Bi directional Long Short-Term Memory (AWBiLSTM) model for precise crop prediction. Given the increasing complexity of DL models and the risk of overfitting, effective feature selection becomes essential. Our MSSA-AWBiLSTM approach uses real-time soil and climate data to provide fast, accurate crop recommendations, enabling informed decision-making and improved yield outcomes.

2 Literature Review

Soni et al. [25] used ensemble ML but lacked deep temporal modeling and adaptive optimization for crop recommendation based on soil parameters using XGBoost and decision tree. In [24] the potential of LSTM networks is emphasized, but the lack of integration with optimization techniques is noted for the prediction and recommendation of crop yields. By combining an adaptive weighting-based ensemble of BiLSTM networks with a MSSA, this research presents a novel hybrid framework for crop recommendation. The suggested method makes use of MSSA for dynamic hyperparameter optimization and model selection, which improves convergence efficiency and prevents local optima in contrast to current methods that rely on static machine learning

models or isolated deep networks. Furthermore, adaptive weighting is used by the ensemble method to highlight the predictions of more dependable BiLSTM learners, enhancing the system's overall generalization and robustness. The model performs noticeably better than state-of-the-art methods thanks to this synergistic integration, which allows it to capture intricate temporal correlations in agricultural data while adjusting to changing environmental and soil conditions.

For agriculture to be sustainable and to maximize output, accurate crop recommendations are essential. Single deep learning methods frequently lack robustness, while traditional machine learning models have trouble with nonlinear, temporal agricultural data. This paper suggests a novel hybrid technique that combines an adaptive weighting-based ensemble of bidirectional LSTM networks with a MSSA for improved hyperparameter tuning in order to overcome these constraints. This integration increases prediction accuracy, speeds up convergence, and captures bidirectional temporal trends. For farmers and agricultural planners, the suggested system provides a scalable and intelligent decision-support tool that enables more adaptive and knowledgeable crop selection depending on soil and environmental factors.

The primary contributions of the MSSA-AWBiLSTM CRS are outlined:

- Initially, pre-processing is conducted to improve the quality of the real-time soil and climatic parameters.
- MSSA is then utilized for feature selection. The objective is to swiftly identify the most informative and relevant features, ensuring reliable prediction results.
- Crop recommendations are generated using the BiLSTM Network Ensemble via the Adaptive Weighting approach, facilitating farmers in obtaining instantaneous and accurate crop recommendations by inputting their preferred climate and crop attributes.
- To validate the proposed method, extensive comparisons are conducted with other conventional frameworks, focusing on precision, recall, and execution time.

The remaining sections are organized as follows: Section 2 describes the mechanism for the proposed CRS. Section 3 presents the results and discussions.

Finally, Section 4 summarizes the work and suggests directions for further research.

3 Methodology

MSSA with AWBiLSTM is designed for optimal crop prediction and recommendation utilizing optimal feature set chosen according to their fitness values. The primary aim is to guide farmers in suitable crop cultivation based on real-time soil and climate information gathered using IoT based sensors for increased crop productivity and maximizing yield.

3.1 Data collection and Preprocessing

The crop recommendation dataset, available at Kaggle (<https://www.kaggle.com/siddharthss/crop-recommendation-dataset/>), contains vital agricultural parameters such as rainfall, crop yield, nitrogen (N), phosphorus (P), potassium (K) levels, temperature, humidity, pH, and precipitation. To ensure high-quality inputs, data preprocessing is performed—addressing missing values (denoted by '.'), removing duplicates, and standardizing the structure. Missing entries are replaced with distinct large negative values to help models identify them as outliers. As the dataset lacks class labels, supervised learning labels are generated based on crop yield and cultivated area: crops with an area > 0 are labeled as class 1, otherwise class 0.

Sensors are used to monitor soil moisture, pH, temperature, humidity, and rainfall, which are essential for accurate crop recommendations. Additionally, mobile cameras detect crop infections to suggest suitable fertilizers, with all data securely stored in the cloud. The dataset is split into training and testing sets, and deep learning (DL) algorithms are used to train the crop recommendation model effectively.

3.2 MSSA algorithm for feature selection

In this proposed modification of the SSA (MSSA) [31], the concept of the local best solution is introduced. The mathematical model of the MSSA is kept unchanged compared to SSA for its leader salp position update, while the followers salp positions are updated using the following equation:

$$x_j^i = b_j^i + \epsilon(x_j^q - x_j^i), j = 1, \dots, D \quad (1)$$

where b_j^i is the i^{th} local best position in the j^{th} dimension, x_j^i is the i^{th} position in the j^{th} dimension, x_j^q is a randomly selected position in a j^{th} dimension

from a set of N solution. dimension. While the best fitness values B_i are assigned the corresponding outcomes F_i 's, the best local solutions b_i 's are first made equal to the first created population. As the least (or greatest) fitness function value F_i and its accompanying solution x_i , the global optimum G and its corresponding solution g are chosen.

Inside the main loop, Eq. (1) is used to update the leader's salp position in [31], and Eq. (1) is used to calculate the new follower salp's locations. The newly created positions' feasibility is confirmed, and the associated fitness function is assessed and stored. By comparing the recently acquired best local optimums to the previously stored global optimum, the MSSA's final step finds a new global optimum. Once a predetermined maximum number of generations has been achieved, the iteration ends. In contrast to SSA, MSSA incorporates information on the local best in each individual of the population into its exploration techniques. This could be useful if the food supply becomes trapped in a local minimum and enable the finding of other food sources using the modified MSSA method. However, each neighborhood of the local best is additionally exploited using a step size that is established by comparing the associated local position x^i with a random individual x^q . This enables the neighborhood search area surrounding the local best to be expanded or reduced.

Each solution is represented by a binary value ("1" denoting a feature that was selected and "0" denoting features that were not), and a random population of salps is produced. The food fitness of each solution (climate feature) is then calculated by MSSA. Next, the food source is selected from the solution with the lowest fitness value. As a result, it uses the K-Nearest Neighbor (KNN) based fitness goal function to choose the best answer. In the subsequent phase, MSSA uses Eq. (1) in [31] to modify the salp positions. Using climate data, the fitness value is updated at each cycle to update the optimal solution and enhance the relevant feature. The optimal solution is provided as a vector of binary values once the maximum number of iterations has been reached. On the dataset's testing segment, KNN makes use of the chosen features. It is used in conjunction with the KNN classifier to assess it on feature selection problems. Therefore, MSSA is used to identify feature combinations that use the fewest features possible while maximizing classification precision. Using training data, the fitness function is changed to improve crop prediction performance on the validation dataset.

Therefore, using climatic data as a reference, the salp with the highest fitness score is modified to improve the chosen features. Salps with the highest total fitness score related to food discovery process are superior, and optimal values for each salp (matching to climate features) are established. The MSSA method uses a small set of attributes to find feature combinations that maximize classification accuracy. In the feature space, every feature has a distinct dimension that ranges from 0 to 1.

3.3 Long Short Term Memory Network

A specific kind of Recurrent Neural Network called LSTM was created to solve the vanishing gradient problem and learn long-term dependencies. Every LSTM cell has a hidden state (h_t) for short-term output and a cell state (C_t) for long-term memory. Three gates are used, namely:

The forget gate determines which data from the prior cell state should be discarded.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

The input gate determines which additional data should be added to the cell state.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad \tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (3)$$

The input gate decides what portion of the cell state is output.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (4)$$

Cell state Update equation:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

Hidden State (Output):

$$h_t = o_t * \tanh(C_t) \quad (6)$$

Each gate controls the cell state using a tanh function and permits or prohibits information flow using a sigmoid function. Retained old memory and newly acquired information are combined to update the cell state.

3.4 Crop recommendation using BiLSTM Network Ensemble via Adaptive Weighting method

This research employs a BiLSTM Network Ensemble via Adaptive Weighting to predict optimal crop conditions by learning from historical data. The system enables farmers to quickly receive accurate crop recommendations based on climate and crop preferences. LSTM networks are well-suited for handling sequential data due to their ability to capture long-term dependencies through memory cells and gated structures. These networks efficiently extract high-level temporal features using recurrent hidden layers, making them effective in representing evolving input sequences [21]. LSTMs operate over time-series data by using a sequence length parameter, which reflects changes in input vectors across time steps. Each memory block includes a memory cell, input gate, forget gate, and output gate. The memory cell retains temporal knowledge, while the gates manage the flow of information. At each time step t , the LSTM updates the memory cell c_t and generates a hidden state h_t . These operations are governed by a set of equations that define the internal dynamics of modern LSTM models with forget gates [22].

$$i_t = \sigma(W_{xi}x_t + W_{hi}x_{t-1} + b_i) \quad (7)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}x_{t-1} + b_f) \quad (8)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}x_{t-1} + b_o) \quad (9)$$

$$c_t = f_t \oplus c_{t-1} + i_t \oplus \phi(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (10)$$

$$h_t = o_t \oplus \phi(c_t) \quad (11)$$

where W_{mn} represents the weight of the link between gate m and n , and b_n acts as a bias parameter to be learned, where $m \in x, h$ and $n \in i, f, o, c$. Furthermore, \oplus represents the Hadamard product, σ designates the normal logistic sigmoid function, and ϕ represents the hyperbolic tangent function. $\sigma(x) = 1/(1 + e^{-x})$; $\phi(x) = (e^x - e^{-x})/(e^x + e^{-x})$. The input, forget, and output gates are indicated as i_t , f_t , and o_t , respectively; c_t represents the internal state of the memory cell c at time t . It's hidden layer at time t is represented by the vector h_t , whereas h_{t-1} represents the values output by

each memory cell in the hidden layer at the previous time step.

This paper proposes an innovative BiLSTM ensemble recommendation approach to improve prediction accuracy. The method combines outputs from multiple BiLSTM models by dynamically adjusting their weights based on previous prediction errors and a forgetting factor. These adaptive weights enhance the ability of the system to generate accurate recommendations. Each LSTM network in the ensemble serves as a base predictor [23]. To promote diversity and capture complex data patterns, individual networks are trained with varying sequence lengths and memory cell configurations. This variation allows the ensemble to better model non-linear dependencies and complement each other's strengths, resulting in more precise and robust predictions.

Assuming a total of M LSTM models, their combined prediction for the time series, denoted as $(y^{(1)}, y^{(2)}, \dots, y^{(N)})$ with N observations, can be shown as:

$$\hat{y}^k = \sum_{m=1}^M w_m \hat{y}_m^k, \quad \text{for } k = 1, \dots, N \quad (12)$$

The recommendation output (at the k th time stamp) derived using m th LSTM model denoted by \hat{y}_m^k , while the related combining weight is w_m . Each weight w_m corresponds to certain model's output. We have $0 \leq w_m \leq 1$ and $\sum_{m=1}^M w_m = 1$.

BiLSTM is an advanced form of LSTM that processes input sequences in both forward and backward directions. This dual processing allows BiLSTMs to capture contextual information from both past and future elements, making them highly effective for sequence modeling tasks. A BiLSTM consists of two LSTM layers: one reads the input from start to end, while the other reads from end to start. Their outputs are combined at each time step, offering a richer understanding of the sequence. BiLSTMs excel in capturing long-range dependencies and are especially useful in tasks like time series prediction, natural language processing, and speech recognition, where full context is essential for accurate predictions.

3.4.1 Adaptive weights for BiLSTM models

A distinctive weight determination method is used to accommodate the changing dynamics of the underlying time series data in a flexible manner. When the regression errors of individual estimators

exhibit zero mean and are uncorrelated, their weighted aggregation achieves minimal variance by assigning weights inversely proportional to the variances of the estimators [26]. This principle forms the theoretical foundation of our weight determination solution. A novel strategy for weight determination is utilized, with the goal of dynamically capturing the evolving dynamics of the underlying time series. The computation of the combining weights is conducted recursively:

$$w_m^{k=1} = w_m^k + \lambda \Delta w_m^k, \quad \text{for } m = 1, \dots, M \quad (13)$$

where $\lambda = 0.3$, and Δw_m^k is computed according to the inverse prediction error of the respective BiLSTM base model:

$$\Delta w_m^k = \frac{\frac{1}{\epsilon_m^k}}{\frac{1}{\epsilon_m^1 + \epsilon_m^2 + \epsilon_m^3 + \dots + \epsilon_m^k}} \quad (14)$$

The ϵ_m^k is related to past prediction error measured up to the k^{th} time step:

$$\epsilon_m^k = \sum_{t=k-V+1}^k \gamma^{k-t} e_m^t \quad (15)$$

where $0 < \gamma \leq 1$, $1 \leq V \leq k$, and e_m^t denotes the prediction error at each time step t of the m th LSTM model, ϵ_m^k is computed by constructing a sliding time window comprising the last V prediction errors. The introduction of the forgetting factor γ serves to reduce the influence of older prediction errors. Through updating weights over time, we can determine these weights by examining inherent patterns in consecutive data forecasting attempts. This approach is suitable as it circumvents the need for complex optimization to ascertain adaptive weights.

Using the preceding Eq., the weights evolve over the time series encompassing T time steps (where $K = 1, \dots, T$), culminating in the ultimate weights w_m^T (where $m = 1, \dots, M$). Eventually, the weight designated to each model is calculated as:

$$w_m = \frac{w_m^T}{\sum_{n=1}^M w_n^T} \text{ for } m = 1, \dots, M \quad (16)$$

The weights computed satisfy the constraints such that $0 \leq w_m \leq 1$ and $\sum_{m=1}^M w_m = 1$ [27].

The equations presented indicate that adaptive weight based amplification is effectively carried out within the LSTM. Therefore, it proves instrumental in generating precise crop recommendation outcomes [27]. Figure 1 shows the general architecture of AWBiLSTM model.

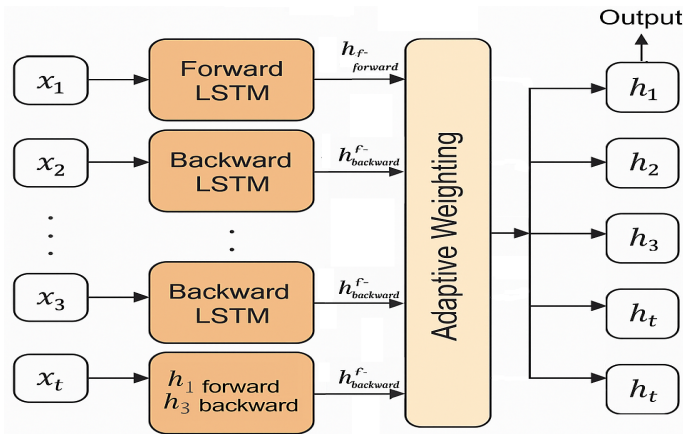


Figure 1. General architecture of AWBiLSTM model.

4 Experimental Results and Discussion

The performance evaluation of the MSSA-AWBiLSTM method is carried out using Python 3.6.5 on a PC equipped with an i5-8600k processor, GeForce 1050Ti 4 GB GPU, 16 GB RAM, 250 GB SSD, and 1 TB HDD. The proposed MSSA-AWBiLSTM is compared with existing AWBiLSTM model to assess its effectiveness. Notably, the proposed MSSA-AWBiLSTM model surpasses existing classification algorithms [18, 19, 28–30] for crop recommendation across various performance metrics including recall, precision, accuracy, specificity, Precision Recall Curve (PR-Score), receiver operating characteristic (ROC-Score), F1-Score, Matthews correlation coefficient (MCC), R2-Score and execution time. These algorithms have previously demonstrated their performance in related literature. The comparison of performance evaluation metrics is presented in Table 1.

Table 1 shows the crop recommendation results obtained from the MSSA-AWBiLSTM. Results demonstrate superior performance with an accuracy of 98.72%, precision of 98.81%, recall of 98.54%, specificity of 98.10%, PR-score of 98.18%, ROC-score of 98.49%, F1-score of 98.48%, and MCC of 97.63%.

Table 2 presents a comparative analysis of crop recommendation results achieved by the MSSA-AWBiLSTM model alongside existing models,

Table 1. Performance evaluation of MSSA-AWBiLSTM across standard classification metrics.

Performance Metrics	Values
Accuracy	98.72
Precision	98.81
Recall	98.54
Specificity	98.10
PR-Score	98.18
ROC-Score	98.49
F1-Score	98.48
MCC	97.63

Table 2. Comparative analysis of MSSA-AWBiLSTM and benchmark models on classification performance.

Methods	Accuracy	Precision	Recall	F1-Score
XAI-CROP	95.12	95.62	95.78	95.86
IDCSO-WLSTM	92.68	90.88	91.98	91.79
MSSA-AWBiLSTM	98.72	98.81	98.54	98.48
RFOERNN	98.45	98.51	98.45	98.46
MMML	97.91	97.97	97.91	97.92
NC-SAE	94.64	94.06	94.78	95.49
SVM-Kernel	91.73	91.13	92.42	93.80
SVM	89.49	88.17	88.70	88.46
SSAE-CNN	90.85	93.86	90.60	92.94
PCA-CNN	88.62	89.27	87.75	89.08
DT	85.07	84.73	85.91	85.39

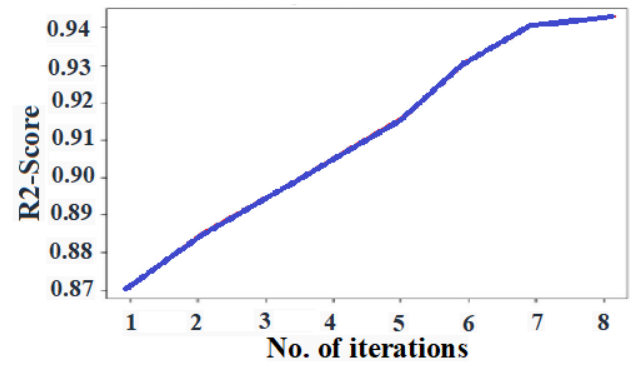
including [12, 19, 28–30]. The crop recommendation outcomes of the MSSA-AWBiLSTM model are compared with those of existing models in terms of R2 score, as presented in Table 3. These values affirm the superior crop recommendation capabilities of the MSSA-AWBiLSTM model compared to other models.

When the proposed method operates at a faster pace, it signifies an improvement in the system's efficiency. Both the proposed MSSA-AWBiLSTM algorithm and the current methods are utilized to compare their execution times as a measure of comparison. Remarkably, the MSSA-AWBiLSTM algorithm demonstrates shorter execution times when applied to the crop recommendation dataset. Table 4 provides a comparison between the existing and proposed systems based on Execution Time (in seconds).

Figure 2 depicts the R2 convergence curve derived from the MSSA-AWBiLSTM model. This curve gives

Table 3. Comparison of R2 score analysis.

Methods	R2-Score
XAI-CROP	94.15
IDCSO-WLSTM	92.54
MSSA-AWBiLSTM	99.14
RFOERNN	99.88
MMML	98.54
SVR	91.99
KNN	87.05
MLR	89.10
ANN	91.97

**Figure 2.** R2 convergence curve for MSSA-AWBiLSTM model.**Table 4.** Comparison of Execution time of MSSA-AWBiLSTM approach with existing methodologies.

Methods	Execution time (sec)
XAI-CROP	240.0554
IDCSO-WLSTM	241.0484
MSSA-AWBiLSTM	228.0112
RFOERNN	240.0529
MMML	242.0492
SVR	242.0532
KNN	241.0639
MLR	242.0677
ANN	242.0704

crucial information about the model's performance and capacity to simulate the dataset's intrinsic patterns and connections. The R2 convergence curve, which serves as an indicator of convergence during training, shows the R2 score on the y-axis against the number of iterations on the x-axis. The R2 score indicates how much of the target variable's variance is explained by the proposed model. A higher R2 value shows that the model is more aligned with the data and has enhanced prediction accuracy. Analysis of the R2 convergence curves allows researchers to gauge training dynamics, performance stability, convergence rate, and the potential for overfitting or underfitting.

4.1 Multivariable Analysis

Soil pH, nitrogen (N), phosphorus (P), potassium (K), organic carbon, temperature, rainfall, humidity, sunshine hours, soil type index, and area ID were among the important variables that were included in the multivariable analysis for crop recommendation. Principal component analysis (PCA) was performed to reduce dimensionality and maintain 95% of the variance using the top six principal components, while multiple linear regression (MLR) was utilized to evaluate the effect of these factors on crop

yield. Additionally, the analysis used the variance inflation factor (VIF) to check for multicollinearity and the analysis of variance (ANOVA) to assess the significance of the grouped variables. The AWBiLSTM ensemble model was used to improve prediction after the MSSA selected features based on the PCA results. The main conclusion of the analysis were that temperature, rainfall, and soil pH all significantly affected crop yield ($p < 0.01$). Furthermore, when multivariable features were used instead of raw input data, AWBiLSTM performance increased.

5 Conclusion and Future Work

Agriculture is the primary livelihood for farmers, making crop selection based on soil conditions crucial for maximizing yield. Data-driven recommendations using ML and IoT can significantly improve farmers' decision-making, reducing costs and enhancing precision in precision agriculture. This study presents a novel MSSA-AWBiLSTM method for crop recommendation and yield prediction, addressing the challenge of selecting the most suitable crop during the growing season. The model is fine-tuned using region-specific data and validated for different soil types.

The MSSA algorithm selects key features from the dataset, while AWBiLSTM predicts and recommends crops using an LSTM network ensemble with adaptive weighting. Experimental results show that MSSA-AWBiLSTM outperforms other methods, achieving a 98.72% improvement in crop prediction accuracy and a higher R2 value by optimally selecting climate parameters.

While the MSSA-AWBiLSTM method improves crop prediction, it may not perform well with weed species detection. Future work will focus on expanding the dataset with more variables, integrating

yield prediction, and incorporating market factors, such as post-harvest storage and profitability, to further enhance and validate the proposed model's applicability.

The MSSA-AWBiLSTM approach has drawbacks despite its great accuracy. It might have trouble with tasks that need for multiple kinds of data, including inputs based on images, like weed species detection. Additionally, the model's scalability may be constrained by the availability and quality of region-specific data. Furthermore, outside variables like insect outbreaks and market dynamics are not taken into account. Future studies will concentrate on adding more economic and environmental factors, broadening the dataset to encompass a variety of crop and soil types, and improving the model to facilitate post-harvest decision-making, storage planning, and profitability forecasting for more useful implementation.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The author declare no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

- [1] Zhai, Z., Martínez, J. F., Beltran, V., & Martínez, N. L. (2020). Decision support systems for agriculture 4.0: Survey and challenges. *Computers and Electronics in Agriculture*, 170, 105256. [CrossRef]
- [2] Bondre, D. A., & Mahagaonkar, S. (2019). Prediction of crop yield and fertilizer recommendation using machine learning algorithms. *International Journal of Engineering Applied Sciences and Technology*, 4(5), 371-376.
- [3] Akulwar, P. (2020). A recommended system for crop disease detection and yield prediction using machine learning approach. *Recommender System with Machine Learning and Artificial Intelligence: Practical Tools and Applications in Medical, Agricultural and Other Industries*, 141-163. [CrossRef]
- [4] Kumar, A., Sarkar, S., & Pradhan, C. (2019, April). Recommendation system for crop identification and pest control technique in agriculture. In *2019 International Conference on Communication and Signal Processing (ICCSP)* (pp. 0185-0189). IEEE. [CrossRef]
- [5] Shahana, I. L., Kamal, V. P., Anoop, R. L., & Francis, A. (2022). Crop prediction methods: A comparative study. *Asian J Agric Ext Econ Sociol*, 40(8), 228-236.
- [6] Gosai, D., Raval, C., Nayak, R., Jayswal, H., & Patel, A. (2021). Crop recommendation system using machine learning. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(3), 558-569.
- [7] Das, B., Ali, S. M., Shaikh, M. Z., Chandio, A. F., Rahu, M. A., Pabani, J. K., & Khalil, M. U. R. (2023, January). Linear regression based crop suggestive system for local pakistani farmers. In *2023 Global Conference on Wireless and Optical Technologies (GCWOT)* (pp. 1-6). IEEE. [CrossRef]
- [8] Ray, R. K., Das, S. K., & Chakravarty, S. (2022, January). Smart crop recommender system-a machine learning approach. In *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 494-499). IEEE. [CrossRef]
- [9] Ujjainia, S., Gautam, P., & Veenadhari, S. (2021). A crop recommendation system to improve crop productivity using ensemble technique. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 10(4).
- [10] Pudumalar, S., Ramanujam, E., Rajashree, R. H., Kavya, C., Kiruthika, T., & Nisha, J. (2017, January). Crop recommendation system for precision agriculture. In *2016 eighth international conference on advanced computing (ICoAC)* (pp. 32-36). IEEE. [CrossRef]
- [11] Venkatanareesh, M., & Kullayamma, I. (2024). Deep learning based concurrent excited gated recurrent unit for crop recommendation based on soil and climatic conditions. *Multimedia Tools and Applications*, 83(24), 64109-64138. [CrossRef]
- [12] Kiruthika, S., & Karthika, D. (2023). IOT-BASED professional crop recommendation system using a weight-based long-term memory approach. *Measurement: Sensors*, 27, 100722. [CrossRef]
- [13] Tubishat, M., Ja'afar, S., Alswaitti, M., Mirjalili, S., Idris, N., Ismail, M. A., & Omar, M. S. (2021). Dynamic salp swarm algorithm for feature selection. *Expert Systems with Applications*, 164, 113873. [CrossRef]
- [14] Mirjalili, S., Gandomi, A. H., Mirjalili, S. Z., Saremi, S., Faris, H., & Mirjalili, S. M. (2017). Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Advances in engineering software*, 114, 163-191. [CrossRef]
- [15] Faris, H., Heidari, A. A., Ala'M, A. Z., Mafarja, M., Aljarah, I., Eshtay, M., & Mirjalili, S. (2020). Time-varying hierarchical chains of salps with random weight networks for feature selection. *Expert Systems with Applications*, 140, 112898. [CrossRef]

- [16] Kassaymeh, S., Abdullah, S., Al-Betar, M. A., & Alweshah, M. (2022). Salp swarm optimizer for modeling the software fault prediction problem. *Journal of King Saud University-Computer and Information Sciences*, 34(6), 3365-3378. [CrossRef]
- [17] Aljarah, I., Habib, M., Faris, H., Al-Madi, N., Heidari, A. A., Mafarja, M., ... & Mirjalili, S. (2020). A dynamic locality multi-objective salp swarm algorithm for feature selection. *Computers & Industrial Engineering*, 147, 106628. [CrossRef]
- [18] Shams, M. Y., Gamel, S. A., & Talaat, F. M. (2024). Enhancing crop recommendation systems with explainable artificial intelligence: a study on agricultural decision-making. *Neural Computing and Applications*, 36(11), 5695-5714. [CrossRef]
- [19] Gopi, P. S. S., & Karthikeyan, M. (2024). Red fox optimization with ensemble recurrent neural network for crop recommendation and yield prediction model. *Multimedia Tools and Applications*, 83(5), 13159-13179. [CrossRef]
- [20] Colombo-Mendoza, L. O., Paredes-Valverde, M. A., Salas-Zárate, M. D. P., & Valencia-García, R. (2022). Internet of Things-driven data mining for smart crop production prediction in the peasant farming domain. *Applied Sciences*, 12(4), 1940. [CrossRef]
- [21] Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019*.
- [22] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780. [CrossRef]
- [23] Adhikari, R. (2015). A neural network based linear ensemble framework for time series forecasting. *Neurocomputing*, 157, 231-242. [CrossRef]
- [24] Oikonomidis, A., Catal, C., & Kassahun, A. (2023). Deep learning for crop yield prediction: a systematic literature review. *New Zealand Journal of Crop and Horticultural Science*, 51(1), 1-26. [CrossRef]
- [25] Soni, T., Gupta, G., & Dutta, M. (2024, November). A Comparative Analysis of Decision Trees, Random Forests, and XGBoost for Enhanced Crop Recommendation. In *2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS)* (pp. 626-630). IEEE. [CrossRef]
- [26] Perrone, M. P., & Cooper, L. N. (1995). When networks disagree: Ensemble methods for hybrid neural networks. In *How we learn; How we remember: Toward an understanding of brain and neural systems: Selected papers of Leon N Cooper* (pp. 342-358). [CrossRef]
- [27] Choi, J. Y., & Lee, B. (2018). Combining LSTM network ensemble via adaptive weighting for improved time series forecasting. *Mathematical problems in engineering*, 2018(1), 2470171. [CrossRef]
- [28] Dash, R., Dash, D. K., & Biswal, G. C. (2021). Classification of crop based on macronutrients and weather data using machine learning techniques. *Results in Engineering*, 9, 100203. [CrossRef]
- [29] Gopal, P. M., & Bhargavi, R. (2019). A novel approach for efficient crop yield prediction. *Computers and Electronics in Agriculture*, 165, 104968. [CrossRef]
- [30] Zhang, W. T., Wang, M., Guo, J., & Lou, S. T. (2021). Crop classification using MSCDN classifier and sparse auto-encoders with non-negativity constraints for multi-temporal, Quad-Pol SAR data. *Remote Sensing*, 13(14), 2749. [CrossRef]
- [31] Ouaar, F., & Boudjemaa, R. (2021). Modified salp swarm algorithm for global optimisation. *Neural Computing and Applications*, 33(14), 8709-8734. [CrossRef]



Parijata Majumdar is an accomplished academican and researcher in the field of Computer Science and Engineering. She currently serves as an Assistant Professor at the Indian Institute of Information Technology (IIIT), Agartala. Prior joining IIIT Agartala she was an Associate Professor at Techno College of Engineering, Agartala. She completed her Ph.D. in Computer Science and Engineering from the National Institute of Technology (NIT) Agartala in 2023. She holds an M.Tech. degree in Computer Science and Engineering from Tripura University, where she was a gold medalist. She earned her B.E. in Computer Science and Engineering from Tripura Institute of Technology, Narsingarh. Dr. Majumdar's research interests include Artificial Intelligence, Machine Learning, Internet of Things (IoT), Precision Agriculture, Blockchain Technology, Image Processing, and Metaheuristic Algorithms. She has published several papers in SCI and SCOPUS-indexed journals and has contributed to patent filings. She is also an active reviewer and invited speaker for various international journals and conferences. (Email: parijata.majumdar@iiitagartala.ac.in)