Journal of Advances in Management, Engineering and Science (JAMES), July 2026; 1(3): 57-67

Research Article

Hybrid Artificial Intelligence Framework for Soil Analysis and Crop Recommendation in Smart Agriculture

***Manikandan K**, Assistant Professor, Department of Plant Pathology, School of Agricultural Science, Dhanalakshmi Srinivasan University, Samayapuram, Tiruchirappalli, Tamil Nadu,India.*
manipatho23194@gmail.com

Received: 16th May 2026; Accepted: 28nd June 2026; Published: 16th July 2026

https://doi.org/10.65470/james.v1i03.36

***Abstract** – Sustainable agricultural strategies that optimise food production while minimising resource depletion are in high demand due to the increasing global population. Agricultural output and decision-making are both enhanced by Smart Farming's crop suggestion and soil analysis, which take into account climatic conditions, environmental considerations, and soil fertility when deciding which crops to plant. This research presents a smart system that uses topography, soil characteristics, climatic variables, and past crop yield data to combine soil analysis with sophisticated machine learning. By standardising, addressing missing values, and encoding categories, data quality and consistency are enhanced. To find the most important traits and get rid of duplicates, a multi-stage feature selection method is used, which includes PCA, IG, and correlation analysis. After that, a hybrid model called Kernel ELM-Attn is used to forecast crop yield. The ST-GETN model is then optimised for accurate crop recommendation using the PantheraCobra metaheuristic. The experimental results demonstrate that the suggested framework outperforms existing methods and proposes crops with a 97.26% accuracy rate. The suggested method improves prediction accuracy, precision farming, soil-based decision-making, and long-term Smart Farming viability via resource optimisation and increased crop yields.*

***Keywords**— Precision Agriculture (PA), Climate Smart Agriculture (CSA), Soil Organic Carbon (SOC), Principal Component Analysis (PCA), Information Gain (IG), Spatio-Temporal Group-Enhanced Transformer Network (ST-GETN), Spatial Group (SG).*

INTRODUCTION

The expansion of arable land cannot occur without farming. About 70% of India's workforce is involved in agriculture, which is also one of the country's most important economic sectors. Problems in agriculture have long impeded the progress of the nation. To overcome this obstacle, we need to use smart farming practices, which include updating traditional farming techniques [1]. In our view, smart farming makes use of the Internet of Things (IoT), smart technology, AI, and automated gadgets. It can be costly and time-consuming to adopt these technologies. One important aspect of farming is PA, which heavily relies on data transmission technologies. The soil's mineral composition can be ascertained through laboratory or sensor-based soil testing. A variety of sensors enable the gathering of data in real-time. In order to improve cropping, we need to identify which crops are best suited to each field. Factors such as water availability, soil temperature, wind speed and direction, pollution levels, soil moisture, and other weather-related variables affect crop prediction and production. Soil parameters such as phosphorus (P), potassium (K), nitrogen (N), humidity, temperature, water pollution, water level, and pH value should be measured area-wise using sensors. With these details, they can suggest crops according to their optimal needs. Nevertheless, these sensors necessitate substantial investment, specialised knowledge, and regular maintenance [2].

To guarantee long-term food security in the face of climate change, proponents of climate-smart agriculture (CSA) argue that farmers should follow a methodical approach to strategy development. Soil carbon sequestration and soil health are two of CSA's primary goals, along with reducing emissions of greenhouse gases. More carbon can be sequestered in soils by choosing to increase carbon inputs and decrease carbon outputs. Soil organic carbon (SOC) sequestration methods often advocate for crop rotations that include cover crops, soil applications of biochar, and conservation tillage practices that reduce soil disturbance. Many observations and data have been collected over the past few decades as a result of the widespread application of these management approaches in the world's most important agricultural regions [3].

The following is the outline of the paper: The second section provides a synopsis of relevant literature on smart agricultural recommendation frameworks. Section III provides a detailed description of the approach and illustrates the conceptual architecture of the proposed framework. The findings and analysis of the experiments are presented in Section IV. Section V concludes with a brief summary of the study's findings.

RELATED WORKS

Machine learning has been the tool of choice for a great deal of recent agricultural research [4]. Soil fertility rate was determined using a variety of classification algorithms, including J48, Random Forest (RF), and Naive Bayes (NB) [5]. In order to measure the yield rate and improve the gain, data mining methods were utilised [6]. Additionally, two methods known as Support Vector Machines (SVM) and Artificial Neural Networks (ANN) have been compared in a study [7]. Support Vector Machine (SVM) outperformed ANN, according to the results. Classification has made use of Fuzzy C-Means Clustering [8]. Soil temperature can also be measured using CANFIS, an alternative system that uses Coactive Neuro-Fuzzy Inference. The model generated the maximum amount of inaccuracy, as demonstrated by the results [9]. Additionally, the SaE-ELM algorithm combines the Extreme Learning Machine (ELM) with the Self-Adaptive Evolutionary (SaE) model [10]. Solar radiation, air pressure, and temperature were the primary variables that the model attempted to simulate [11].

For the purpose of this work, the user's crop recommendation was made using Machine Learning [12] techniques including Support Vector Machine, decision Tree, K Nearest Neighbour, Linear Regression model, Neural Network, Naïve Bayes, and Support Vector Machine, rather than SVM and choice tree classifier [13]. It has exposed users to several algorithms that may be compared. The production value was predicted using a Linear Regression model, which takes into account environmental parameters including temperature, rainfall, and humidity [14]. Using a CNN and K-means model, weeds that are cultivated in proximity to soybeans can be identified [15]. A CNN was employed for weed and soybean classification, while K-means was employed for feature identification in the images[16]. Further evidence that fine-tuning the CNN model can enhance accuracy [17].

Intelligent analytics and deep learning are finding more and more uses in this area, particularly in the areas of smart farm automation, soil analysis, disease diagnosis, and agricultural yield estimation [18]. According to a recent analysis of deep learning in agriculture, current research in the field is making more and more use of CNN-, RNN-, and related neural models for smart farming applications, plant health analysis, and yield prediction [19]. The review also highlighted some of the challenges in this field, including insufficient data, poor regional generalisation, and an absence of interpretability [20].

This research makes a unique contribution by creating a unified architecture that integrates advanced feature selection, prediction, optimisation, explainability, and spatio-temporal deep learning techniques. The framework is designed to predict crop yields and make intelligent crop recommendations. Several methodological changes are introduced by the proposed framework, in contrast to previous studies that concentrate on individual prediction models.

1. This research suggests a model called Hybrid Kernel ELM with Attn that uses a combination of nonlinear kernel mapping and an AM to dynamically prioritise agricultural characteristics such soil nutrients, weather, and topography. Compared to traditional ML algorithms, this allows for more precise prediction of agricultural yields.
2. An updated ST-GETN has been developed to better record the interplay between geographical environmental factors and agricultural trends throughout time. The design successfully represents spatial dependencies and sequential temporal fluctuations in agricultural data by merging CNN, Trans Encoder, and BiGRU layers.

METHODLOGY

When it comes to soil nutrient monitoring, "smart farming" means using data-driven methodologies and cutting-edge technology to evaluate and control soil health better. Soil nutrient, pH levels, and moisture can be monitored in real-time by farmers with the use of sensor networks, data analytics, and remote sensing. With this information, they may optimise yield while minimising environmental effect by making informed decisions about fertilisation, irrigation, and crop selection. Sustainable agriculture relies on smart farming for soil nutrient monitoring, which helps farmers become more productive while using less resources. This, in turn, makes farming more resilient and less harmful to the environment.

The data gathering and preprocessing stages are shown in Figure 1, which demonstrates how the proposed model efficiently handles unseen data [21].

The current dataset for India was supplemented with data on precipitation, weather, and fertiliser [22]. Gathered over the period by ICFA, India.

Data fields

N – Soil content Nitrogen ratio

humidity - % in relative humidity

K – Soil content potassium ratio

P – Soil content Phosphorous ratio

temperature - degree Celsius in temperature

PH – Soil PH Value

rainfall – mm Rainfall

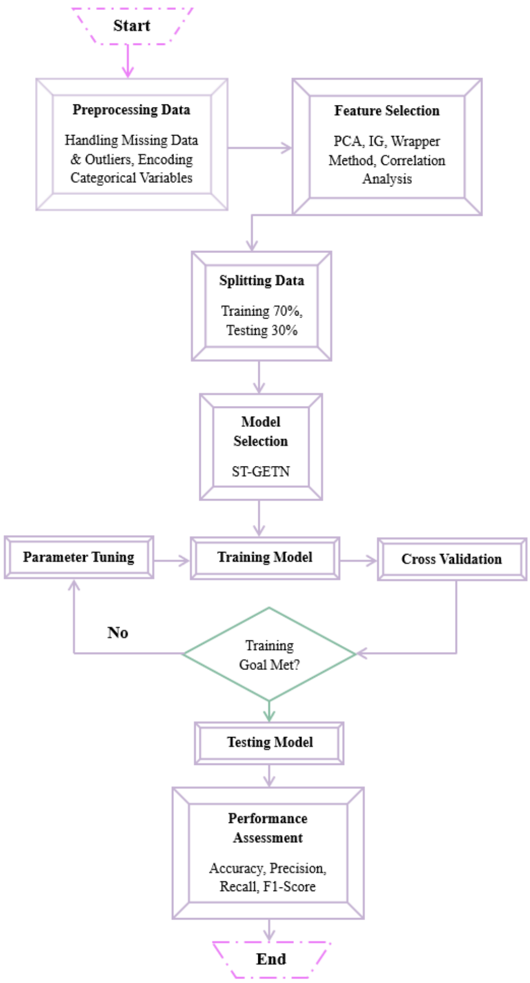

Fig. 1. The Proposed Model Workflow for Smart Agriculture

A. Data Preprocessing:

In the first step, that gather all the data that can find. This includes soil nutrient profiles, measurements from environmental sensors in real time, weather records going back decades, and records of crop yields. Normalisation, encoding and missing value handling, approaches are all part of data pretreatment that helps make sure DL models work. The intricate web of relationships within the agricultural environment can be better understood with the help of this multi-modal information. The following procedures are involved in data pre-processing:

- Data cleansing is the initial stage and entails finding and fixing incomplete or missing values. Missing entries are typical in real-world agricultural data because of sensor malfunction or irregular recordings. In order to fill these gaps, statistical procedures like median or mean imputation are used, which vary based on the type of feature. Temperature, moisture, and nutrition levels are just a few examples of numerical quantities that are normalised to a regular range, usually between 0 and 1. This checks that the size of any one characteristic doesn't make it the deciding factor in model learning.
- The reduction of outliers is another critical step. Using visual analysis techniques or standard deviation criteria, we can find outliers in the dataset that might be due to inaccurate sensor readings or rare environmental circumstances. The goal is to eliminate or, failing that, rectify these outliers such that they do not affect the model's predictions.
- The collection also contains categorical information, including crop labels (e.g., rice, wheat, maize), which are converted to numerical format using encoding techniques. The data is prepared for use with ML algorithms that use numerical input by means of label encoding or one-hot encoding.

B. Feature Selection:

One of the most important preprocessing steps in machine learning is feature selection, which fixes problems caused by high dimensionality. Utilising a learning algorithm begins with selecting a subset of features from the accessible data. By applying an evaluation criterion to the original, comprehensive feature set, feature selection extracts the most important characteristics and discards the rest, leaving only the features that are useful to the ML model or are statistically significant.

To identify and recommend the optimal crop, one can choose from a variety of characteristics, including N, P, K, humidity, temperature, rainfall, and pH. Each crop's name serves as the experiment's dependent variable [23]. Our suggested approach takes variables that are not dependent on rainfall into account, such as N, P, K, humidity, temperature, pH, and so on. At this stage, the dataset is subjected to a number of feature selection procedures, including filter methods and wrapper methods, such as IG, PCA, and correlation analysis.

In order to find the best indications for the farming system, both approaches were utilised. Filter methods depend on dataset features to determine their significance by evaluating and selecting features independently of any learning process. For each given ML method, the wrapper method considers all possible feature combinations and selects the one that yields the best results. With wrapper techniques, the learning process is optimised by selecting fewer features.

Therefore, there is little room for these ideas to be applied in a broader context. Data that is redundant or noisy can have its most important details extracted using PCA, a straightforward nonparametric technique. The PCA employs an orthogonal transformation to separate linearly uncorrelated features from samples of correlated data. When looking for feature subsets utilising correlation-based FS, the degree of feature redundancy is taken into account.

Finding groups of features with low intercorrelation and high individual correlation with the class is the goal of the assessment method. Instead, IG determines whether a feature is present or absent by calculating the difference in entropy. The challenge of finding a feature's importance in a feature space is tackled here by utilising more general methods, including measuring informational entropy.

TABLE I. SELECTED FEATURES NUMBERS

Feature Selection Algorithm	**Selected Features Count**	**Selected Features**
PCA	6	Temperature, P, pH, humidity, k, N
IG	7	Temperature, N, pH, P, humidity, k, rainfall
Correlation	7	Temperature, pH, P, k, humidity, N, rainfall.
Wrapper	5	N, K, P, Rainfall, Humidity

Furthermore, filtering algorithms assign a score to each feature before choosing the features with the highest scores. It displays the degree to which each feature agrees or disagrees with the output labels. A wrapper selection technique is used to choose the effective environmental indicators in this study. Table 1 shows how the wrapper feature selection strategy compares to the other feature selection procedures.

C. Model Training:

1) Hybrid Kernel ELM with Attn using Crop Yield Prediction:

Following FS, a reduced feature matrix is formed by retaining just the most useful and discriminative properties. This increases forecast accuracy by reducing redundant and irrelevant information. In order to forecast crop yields, these features are fed into a hybrid kernel ELM that also incorporates an AM. In order to effectively capture nonlinear interactions, the ELM uses a kernel mapping to a feature space with high dimension and projects the inputs through a nonlinear layer that is hidden [24]. By adjusting the weights of the most important features, the AM improves performance even more and allows the model to zero in on crucial yield-related variables. Last but not least, in order to avoid overfitting and guarantee strong accuracy, crop output is predicted by minimising a regularised MSE.

ELM Input: An ELM takes as input the features selected in the first step, which are represented as a matrix V of P samples and n features. The actual crop yield on each sample is measured by the target vector z.

Representation of Hidden Layer: A nonlinear activation function $\varphi(\cdot)$, multiplied by randomly generated weights U and a bias vector d, will convert the input features V in the hidden layer. This process is carried out by the ELM.

Kernel ELM: By utilising a kernel matrix I—which stores the inner products of hidden layer representations—the input characteristics in the kernel-based ELM are implicitly transformed to a high dimensional space. A regularisation term, β, is applied to the output weights λ computed via regulated least squares to prevent overfitting.

AM: Through the computation of attention ratings that captivate their contribution to the prediction, the AM adaptively provides adaptive relevance to the individual features.

2) Proposed ST-GETN Crop Recommendation:

To choose the best crop, the suggested ST-GETN recommendation system takes into account the expected crop yield z in addition to improved soil and environmental factors. Bidirectional and long-term temporal dependencies are captured by using Trans Encoders and BiGRU networks after the combined input is processed through convolutional layers, which extract spatial and contextual information. The most important traits are highlighted by a SG-Enhanced AM, and the best crop is predicted using a Softmax classifier using probability. To ensure transparent, accurate, and data-driven crop selection, explainability techniques like Grad-CAM and Integrated Gradients are used to emphasise important aspects influencing the suggestion.

$$V_{crop} = [V_{soil}, V_{env}, \hat{z}] \in T^{P\times s} \quad (1)$$

V_{crop} is the input matrix for crop recommendation in Equation (1). $V_{soil} \in T^{P\times q}$- stands for optimised soil features. $V_{env} \in T^{P\times g}$- represents environmental features such as temperature, rainfall, and humidity. $\hat{z}$ is the predicted crop yield from the hybrid kernel-based ELM with attention. $s = q + g + 1$_, which is the total number of recommendation features.

a) Convolutional Layers using Feature Extraction

In order to extract hierarchical spatial and contextual patterns,V_{crop}—representing the crop recommendation input—passes forecasted soil, yield, and environmental data through multiple convolutional layers. Layers of convolution extract increasingly complex representations using learnable biases and filters; feature maps encode crucial global and local information. Equations $(2) - (4)$ provide a detailed and organised depiction of the input, which are feature maps.

$$N_1 = Conv(V_{crop}, J_1) + d_1 \quad (2)$$

$$N_2 = Conv(N_1, J_2) + d_2 \quad (3)$$

$$N_3 = Conv(N_2, J_3) + d_3 \quad (4)$$

Layer k contains the output feature map N_k, the kernel of convolution or filter J_k, and the bias term d_k.

b) Temporal Modelling for Dual-Directional

Before learning the past and future temporal dependence, the feature maps concatenate are processed by a Trans Encoder and a BiGRU. Finding forward and backward sequences in BiGRU models, as well as global links across contexts, can be done using Trans Encoder. With the help of this bidirectional processing, the model is able to portray temporal features on a global scale, which encodes dynamic relations across time and improves its predictive power in relation to context.

$$F^{forward}, F^{backward} = BiGRU\big(TranEncoder\ (N_3)\big) \quad (5)$$

$$D_{temporal} = [F^{forward}, F^{backward}] \in T^{P\times 2f} \quad (6)$$

In equations (5) and (6) , it sees that $F^{forward}$ and $F^{backward}$ represent the hidden states from BiGRU, f stands for the hidden size of BiGRU, and $D_{temporal}$ represents the dual-directional temporal features.

c) SG-Enhanced Attn

To improve the temporal feature representation with the most useful and relevant features, the SG-Enhanced Attn method is employed. To capture patterns across different resolutions, it builds a multi-scale representation of data. Then, it uses spatial group wise selection to emphasise important feature sets and minimise irrelevant ones. An attention-weighted feature set is produced by this technique, which improves the model's interpretability and predictive performance by highlighting the key qualities for appropriate crop recommendation. Equation $(7) - (9)$ emphasises key aspects of $D_{temporal}$ through the SG-Enhanced Attn mechanism:

$$QI = QIPGR\big(F_{temporal}\big) \in T^{P\times b} \quad (7)$$

$$F_{SG} = SG(QI) \in T^{P\times b} \quad (8)$$

$$F_{SG-Enhanced\ Attn} = F_{SG}\,(30) \quad (9)$$

In this context, QI implies multi-scale feature representation, $SG(\cdot)$ means spatial group-wise selection function, $F\_(SG - Enhanced\ Attn)$ means refined attention-weighted features, and b means dimensionality after attention.

d) Analysis of Long-Term Temporal

In the long-term temporal analysis step, which follows the attention-refined characteristics of SG-Enhance Attn, a BiGRU succeeds a Trans Encoder. While Trans establishes connections between all input features' global dependencies and relationships, BiGRU takes into account patterns across time, allowing it to effectively incorporate information from both the past and the future. The result is a comprehensive depiction of temporal features that incorporates long-term dependencies, providing a firm basis for effective crop recommendation in Equations $(10) - (11)$.

$$F_{long} = TranEncoder\ (F_{SG-Enhanced\ Attn}) \quad (10)$$

$$F_{final} = BiGRU\ \big(F_{long}\big) \in Q^{P\times f_h} \quad (11)$$

This is where F_{long} denotes the global dependencies picked up by the Trans, F_{final} denotes the final temporal feature representation, and f_h denotes the hidden size following BiGRU.

e) Layer of Cross Classification

The crop classification layer takes the most recent temporal feature and uses it to feed a Softmax classifier, which then determines which crops are most suitable. To determine the likelihood of each crop type, the classifier passes the total number of logits it has computed to the Softmax algorithm. This is how the model determines the likelihood that each crop will be successful; picking the right crop depends on factors including soil type, environmental factors, and expected yield.

$$y_k = U_a F_{final} + d_a, k = 1,2,\ldots,A \quad (12)$$

$$\hat{z}_k = \frac{\exp(y_k - y_{max})}{\sum_{l=1}^{A} \exp(y_k - y_{max})} \quad (13)$$

The variables A, y_k, U_a, d_a represent the classifier's weights and bias, and $\hat{z}_k$ stands for the probability of crop class k in equations $(12) - (13)$.

f) Optimizer of PantheraCobra

A hybrid metaheuristic algorithm, PantheraCobra takes its cues from the mutually beneficial characteristics of tigers and cobras. The algorithm is able to skip to uncharted territories and avoid local minima thanks to the Cobra's rapid strike and abrupt attacks, while the Tiger's power, stealth, and strategic pouncing direct local exploitation, honing potential solutions. In order to optimise the hyperparameters X of deep networks like ST-GETN, PCO strikes a good balance between exploration and exploitation. This leads to improvements in performance measures like accuracy and F1-score. Each module in the suggested architecture has its own set of hyperparameters, which are displayed in Table 2.

TABLE II. TABLE FOR HYPERPARAMETER

Module	Hyperparameter	Values/ Setting
Hybrid Kernel ELM	Number of Hidden Neurons	100-500
Layer of Conv	Number of layers	3
Trans Encoder	Number of layers	2
BiGRU	Number of layers	1-2
SG-Enhanced Attn	Number of Groups	4-8
Classifier of Softmax	Number of Classifiers	C (No. of Crops)
Training	Optimizer	Adam

RESULT AND DISCUSSION

In countries where a large portion of the population is employed in farming or closely related occupations, the agricultural sector plays a crucial role in maintaining food supply and propelling economic advancement. It is essential for national resilience and food sovereignty, in addition to being an economic goal, to increase agricultural output in these contexts. Choose the right crop varieties is one of the most important things that can affect agricultural output.

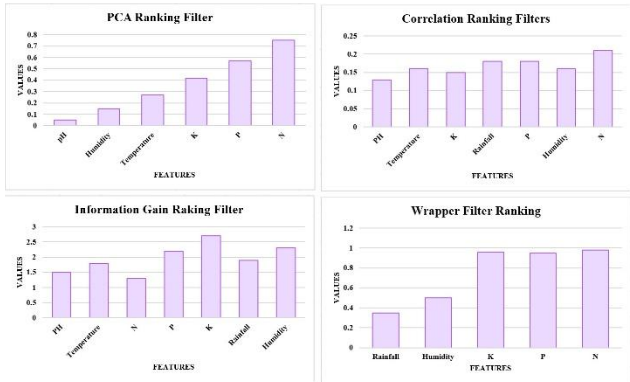

Fig 2. Feature Evaluation

In this investigation, four different types of filters were utilised: IG ranking, correlation, PCA, and wrapper ranking. As shown in Figure 2, these filters were applied to the dataset in order to ascertain which combination of features is more crucial for classification models. To improve the models' performance, we updated the dataset and used the wrapper ranking filter to select fewer attributes. We found that humidity, N, rainfall, P, and K are the five most important attributes. We removed undesirable features, such as temperature and pH, based on the returned ranking of the features.

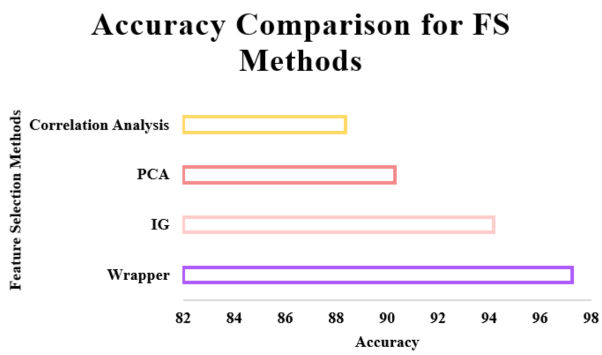

Fig 3. Accuracy Comparison for Feature Selection Methods

Minimal and useful features have been chosen by the wrapper method. The results reveal that wrapper feature selection strategies are more trustworthy than filter methods (Figure 3). This is because wrapper methods only identified the relevant characteristics.

TABLE III. PERFORMANCE COMPARISON (%)

Metric	AM	GNN	XGB	ST-GETN
Accuracy	88.37	90.30	94.18	97.26
Precision	86.45	88.17	92.60	95.54
Recall	84.56	86.49	90.74	93.67
F1-Score	88.84	90.90	94.78	97.95

These outcomes prove that the model is excellent at producing precise ST-GETN and that it has been well-fitted. With an overall accuracy of 97.26% and high precision, recall, and F1-scores across practically all models, the ST-GETN model performs remarkably well, according to the classification report in Table 3.

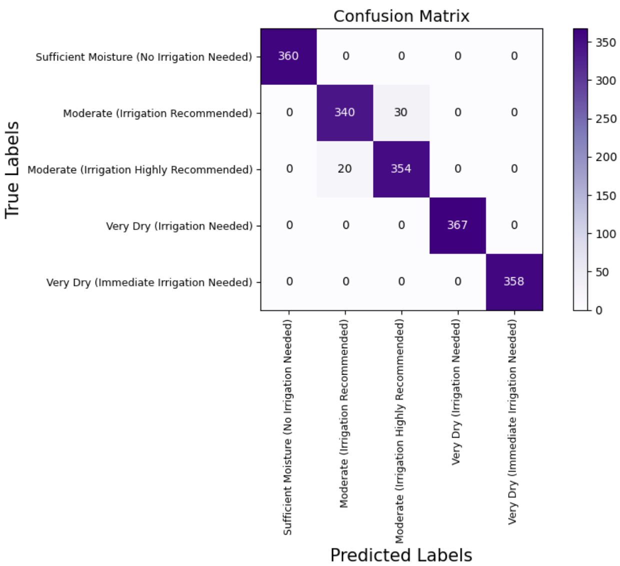

Fig 4. Confusion Matrix

A precision of 97.26% was attained through the application of the ST-GETN model. Figure 4 displays the results of using the confusion matrix to assess how well the ST-GETN model uses soil moisture levels to recommend irrigation. Since the majority of the forecasts matched the observed moisture levels, the data show that the model performed well.

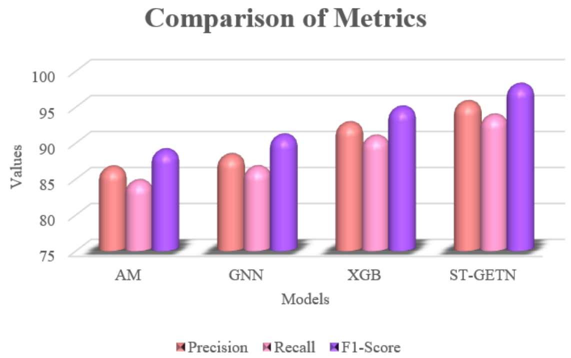

Fig 5. Comparison of Metric

Various models for crop recommendation are compared in Figure 5. The results demonstrate that the ST-GETN model, when implemented, outperforms all others in terms of accuracy (97.26%), precision, recall, and F1-score.

CONCLUSION

The agricultural sector is the backbone of the global economy, contributing 6.4% of GDP and ensuring a steady supply of food. A significant portion of the GDP in nine countries comes from agricultural production. Millions of people rely on the agriculture sector for their energy needs and for the jobs it offers. By 2050, global meat production will have to more than double and wheat production will have to triple to meet human demand. A reliable supply will be ensured by improvements in grain production yields. They need to increase the size of your crop production and embrace a more contemporary approach to farming. The PantheraCobra Optimiser optimises hyperparameters and combines advanced preprocessing with multi-stage feature selection and attention-enhanced deep learning to efficiently capture complex nonlinear relationships in agricultural data with minimal noise and redundancy. Outperforming traditional ML and DL methods, the ST-GETN model provides very trustworthy crop recommendations (accuracy: 97.26%), while the hybrid kernel ELM with Attn produces precise predictions of harvest yields.

Acknowledgement

The author would like to appreciate the effort of the editors and reviewers.

Author Contributions

All authors are equally contributed.

Conflict of Interests

The authors declare that they have no conflicts of interest.

Ethics Approval

There are no human subjects in this article and informed consent is not applicable.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] M. Pyingkodi, K. Thenmozhi, M. Karthikeyan, T. Kalpana, S. Palarimath, and G. B. A. Kumar, "IoT based Soil Nutrients Analysis and Monitoring System for Smart Agriculture," *3rd Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2022 - Proc.*, no. Icesc, pp. 489–494, 2022, doi: 10.1109/ICESC54411.2022.9885371.

[2] M. K. Senapaty, A. Ray, and N. Padhy, "IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture," *Computers*, vol. 12, no. 3, 2023, doi: 10.3390/computers12030061.

[3] X. Bai *et al.*, *Responses of soil carbon sequestration to climate-smart agriculture practices: A meta-analysis*, vol. 25, no. 8. 2019. doi: 10.1111/gcb.14658.

[4] A. Aryanti, N. Iryani, and K. Khairunnisa, "A Smart Recommendation System for Crop Seed Selection Using Gradient Boosting Based on Environmental and Geospatial Data," *J. Appl. Informatics Comput.*, vol. 9, no. 6, pp. 2965–2973, 2025, doi: 10.30871/jaic.v9i6.10249.

[5] S. Ramzan *et al.*, "An innovative artificial neural network model for smart crop prediction using sensory network based soil data," *PeerJ Comput. Sci.*, vol. 10, pp. 1–28, 2024, doi: 10.7717/peerj-cs.2478.

[6] O. Folorunso *et al.*, "GeaGrow: a mobile tool for soil nutrient prediction and fertilizer optimization using artificial neural networks," *Front. Sustain. Food Syst.*, vol. 9, no. March, 2025, doi: 10.3389/fsufs.2025.1533423.

[7] T. Thorat, B. K. Patle, and S. K. Kashyap, "Intelligent insecticide and fertilizer recommendation system based on TPF-CNN for smart farming," *Smart Agric. Technol.*, vol. 3, no. September 2022, p. 100114, 2023, doi: 10.1016/j.atech.2022.100114.

[8] V. Balaji, K. S. Balamurugan, T. Thanarajan, and A. M. Selvi Sundarapandi, "Machine Learning-Based Crop Recommendation System Integrating Soil Properties and Weather Conditions With Iot-Driven Data Collection," *Agrociencia*, vol. 60, no. 1, pp. 138–154, 2026, doi: 10.47163/agrociencia.v60i1.3428.

[9] N. J, V. P. Monickaraj, P. K, and S. K, "PlantIntellect: Smart Crop Recommendation using CNN and Hybrid Weather Analysis," *ITM Web Conf.*, vol. 85, p. 03010, 2026, doi: 10.1051/itmconf/20268503010.

[10] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, and E. H. M. Aggoune, "Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk," *IEEE Access*, vol. 7, pp. 129551–129583, 2019, doi: 10.1109/ACCESS.2019.2932609.

[11] P. Sumathi, V. V. Karthikeyan, M. S. Kavitha, and S. Karthik, "Improved Soil Quality Prediction Model Using Deep Learning for Smart Agriculture Systems," *Comput. Syst. Sci. Eng.*, vol. 45, no. 2, pp. 1545–1559, 2023, doi: 10.32604/csse.2023.027580.

[12] R. Selvanarayanan, S. Rajendran, S. Algburi, O. Ibrahim Khalaf, and H. Hamam, "Empowering coffee farming using counterfactual recommendation based RNN driven IoT integrated soil quality command system," *Sci. Rep.*, vol. 14, no. 1, pp. 1–16, 2024, doi: 10.1038/s41598-024-56954-x.

[13] G. Chauhan and A. Chaudhary, "Crop Recommendation System using Machine Learning Algorithms," *Proc. 2021 10th Int. Conf. Syst. Model. Adv. Res. Trends, SMART 2021*, vol. 3307, pp. 109–112, 2021, doi: 10.1109/SMART52563.2021.9676210.

[14] S. K. Apat, J. Mishra, K. S. Raju, and N. Padhy, "An Artificial Intelligence-based Crop Recommendation System using Machine Learning," *J. Sci. Ind. Res. (India).*, vol. 82, no. 5, pp. 558–567, 2023, doi: 10.56042/jsir.v82i05.1092.

[15] M. Rathi and C. Gomathy, "Smart agriculture

resource allocation and energy optimization using bidirectional long short-term memory with ant colony optimization (Bi-LSTM–ACO),” *Front. Commun. Networks*, vol. 6, no. June, pp. 1–13, 2025, doi: 10.3389/frcmn.2025.1587402.

[16] G. Srivarshini, M. Sujithra, B. Dhanalakshmi, and K. Selvakumar, “Smart farming with agri CNN-LSTM fusion: leveraging soil suitability analysis using real-time sensor data,” *Discov. Appl. Sci.*, vol. 7, no. 10, 2025, doi: 10.1007/s42452-025-07119-6.

[17] S. K. S. Durai and M. D. Shamili, “Smart farming using Machine Learning and Deep Learning techniques,” *Decis. Anal. J.*, vol. 3, no. March, p. 100041, 2022, doi: 10.1016/j.dajour.2022.100041.

[18] A. Filintas, N. Gougoulias, N. Kourgialas, and E. Hatzichristou, “Management Soil Zones, Irrigation, and Fertigation Effects on Yield and Oil Content of Coriandrum sativum L. Using Precision Agriculture with Fuzzy k-Means Clustering,” *Sustain.*, vol. 15, no. 18, 2023, doi: 10.3390/su151813524.

[19] K. P. Swain *et al.*, “Empowering Crop Selection with Ensemble Learning and K-means Clustering: A Modern Agricultural Perspective,” *Open Agric. J.*, vol. 18, no. 1, pp. 1–14, 2024, doi: 10.2174/0118743315291367240207093403.

[20] and P. T. Shriwas, M. K., K. Sindhi, “Deep Learning-Based Multilingual Smart Farming System for Crop Recommendation and Nutrient Monitoring,” *AFRICAN J. Appl. Res.*, vol. 12, no. 5, pp. 260–280, 2026.

[21] A. Khaliq *et al.*, “AI-Driven Smart Agriculture: An Integrated Approach for Soil Analysis, Irrigation, and Crop-Fertilizer Recommendations,” *IEEE Access*, vol. 13, no. July, pp. 141124–141138, 2025, doi: 10.1109/ACCESS.2025.3594162.

[22] “Crop Recommendation Dataset.” Accessed: Jul. 25, 2026. [Online]. Available: https://www.kaggle.com/datasets/siddharthss/crop-recommendation-dataset

[23] D. Garg and M. Alam, “An effective crop recommendation method using machine learning techniques,” *Int. J. Adv. Technol. Eng. Explor.*, vol. 10, no. 102, pp. 498–514, 2023, doi: 10.19101/IJATEE.2022.10100456.

[24] N. Piyar S and D. B. S, “An explainable spatio-temporal deep learning framework for crop yield prediction and recommendation,” *Sci. Rep.*, 2026, doi: 10.1038/s41598-026-54439-7.