International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(12s) 2026 pp. 497-509
https://doi.org/10.70917/ijcisim-2026-3921
Article

Agrodeepfusion: A Hybrid Deep Ensemble Architecture for Crop and Fertilizer Recommendation

Sankareswari K[1], Dr. Sujatha G[2]

[1] PG and Research Department of Computer Science, Sri Meenakshi Govt. Arts College for Women (A), Affiliated to Madurai Kamaraj University, Madurai, India. Assistant Professor, Department of Computer Science, The American College, Madurai, India.
Email: sankari.kaverimanian@gmail.com
[2]PG and Research Department of Computer Science, Sri Meenakshi Govt. Arts College for Women (A), Affiliated to Madurai Kamaraj University, Madurai, India.

Abstract: The rapid advancements in deep learning (DL) methodologies have brought substantial transformations in agricultural practices, particularly in predicting crop and fertilizer recommendation. Accurate prediction of crop is crucial for optimizing food production and addressing food security challenges. While traditional statistical and machine learning approaches have been widely used for years, their limitations in handling large scale, nonlinear and complex agricultural datasets highlight the need for advanced solutions. Also struggle with generalizability and performance under diverse, complex, and resource constrained scenarios. To overcome these challenges, this study proposes AgroDeepFusion, a hybrid deep ensemble framework that integrates Convolutional Neural Network (CNN), Long Short Term Memory (LSTM), and Deep Neural Network (DNN) models. The proposed architecture employs stacking and bagging ensemble techniques to combine spatial feature extraction, sequential dependency modelling, and nonlinear feature learning. A five year soil dataset collected from Indian agricultural soil laboratories in Tamil Nadu containing 22 crop classes and multiple macro and micro nutrient attributes was used for evaluation, Experimental results demonstrate that AgroDeepFusion achieved 95% prediction accuracy, clearly surpassing standalone deep learning models CNN (86.5%), LSTM (87.5%), and DNN (84.0%) when tested separately. Also outperformed existing Machine Learning and Deep Learning baselines in precision, recall, and F1-score. Comparative analysis shows the ensemble approach's effectiveness over standalone models. Furthermore, a rule based fertilizer recommendation module enhances the system's practical usability. This study contributes a novel ensemble architecture and highlights existing gaps in the current literature, offering an effective solution for sustainable agriculture.

Keywords: Crop prediction, Fertilizer recommendation, Deep learning, LSTM, RNN, CNN, Ensemble Technique, Boosting, Stacking, Bagging.

1. Introduction

Predicting crops accurately is crucial for improving agricultural planning and guaranteeing global food security. With the continuous growth in world's population, the demand for efficient and sustainable farming methods has become more pressing than ever [24]. Central to this approach is the accurate prediction of suitable crops for cultivation and the recommendation of appropriate fertilizers to maintain soil fertility. Despite notable advancements in machine learning and deep learning, current crop and fertilizer recommendation systems face notable limitations in terms of adaptability and scalability. Many of these models depend on narrowly focused datasets, which often do not generalize across diverse regions with varying agro climatic conditions. Furthermore, high performance models typically require computational resources that are not readily available in rural or developing agricultural settings. As a result, there is a pressing need for a robust, scalable, and interpretable deep learning framework that can operate effectively on real world soil datasets and provide reliable recommendations even in resource limited environments.

While individual deep learning models have shown promise in agricultural prediction tasks, they often lack robustness and generalizability, particularly when applied to diverse agricultural datasets. Ensemble learning techniques depicted in Figure 1, such as stacking, bagging, and boosting, address these limitations by combining the strengths of multiple models. Stacking leverages outputs from CNNs, LSTMs, and other networks using a meta learner to enhance predictive accuracy [9][13][16] [26], while bagging reduces variance and boosting emphasizes hard to predict samples [20][22][26]. Despite their potential, ensemble methods remain underutilized in crop forecasting, particularly in resource constrained settings where computational efficiency is critical. Previous studies have shown encouraging results using hybrid models like CNN, LSTM, but often fall short in terms of scalability and adaptability across different crops and regions [11][14][23]. This study aims to bridge those gaps by analyzing the scalability, effectiveness and resilience of ensemble deep learning models using real world soil test data. It compares the proposed hybrid approach against existing techniques and emphasizing its practical relevance to precision agriculture.

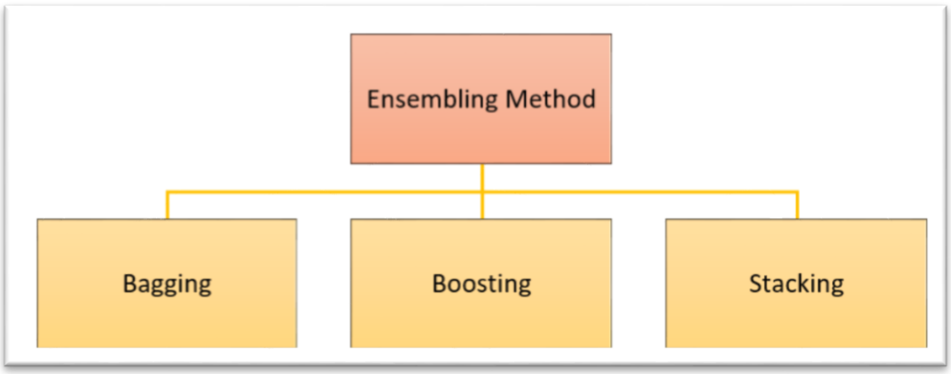

Figure 1: Ensemble Methods

This study proposed AgroDeepFusion ensemble design framework, a hybrid deep learning ensemble model that combines CNN, LSTM, and DNN architectures to recommend optimal crops based on soil profiles. A rule based fertilizer recommendation component, guided by nutrient thresholds, enhances its practical relevance. Using real world five years of soil test data from Indian agricultural labs, the model performance is evaluated and compared with conventional machine learning methods and individual deep learning models. The results show notable enhancements in predictive accuracy and F1-score, while also addressing issues of scalability and computational efficiency. Overall, this work contributes to the development of robust and deployable solutions for precision agriculture.

2. Related Work

In recent years, notable progress has been made in applying machine learning and deep learning techniques to forecast crop yields and recommending fertilizers. For tasks such as crop forecasting and fertilizer recommendations, a range of models has been employed, from traditional statistical techniques like regression to more sophisticated machine learning frameworks, including Support Vector Machines (SVM) and Random Forest [3].

2.1 Traditional Machine Learning based crop recommendation yield prediction

[21] proposed a machine learning model designed to categorize soil samples into four categories such as very high fertile, high fertile, moderately fertile, and low fertile using Support Vector Machine (SVM) techniques. Furthermore, the model recommends fertilizers that could enhance soil fertility and predicts crops suitable for each soil classification determined by the sample group. This model enables farmers to make informed decisions about which crops to cultivate and determine the optimal ratios of Nitrogen (N), Phosphorus (P), and Potassium (K) fertilizers. Compared to K-Nearest Neighbor (KNN), Decision Tree, and other algorithms, SVM demonstrated superior accuracy in its predictions. [6] introduced a farming prediction application designed for agriculturalists. The app uses GPS technology to help farmers to determine their exact location. By leveraging machine learning techniques, it assists in selecting the most profitable crops and forecasting their yields. Several algorithms are utilized for yield prediction, including Support Vector Machine (SVM), Random Forest, Multivariate Linear Regression (MLR), K-Nearest Neighbour (KNN) and Artificial Neural Network (ANN) are employed to predict crop yield. Among these, the accuracy achieved by Random Forest was 95%. Furthermore, the algorithm provides suggestions for fertilizers to maximize crop yields. A model proposed in [7] focuses on location specific crop management, guiding farmers for choosing the right crops based on factors such as planting season, geographic location and soil type.

2.2 Deep Learning based approaches for agricultural prediction

Although these models effectively handle structured information, they encounter difficulties with large and intricate datasets characterized by nonlinear relationships. To further tackle this problem, [17] developed an agro deep learning framework that utilizes deep neural networks to model complex relationships between soil and environmental variables for improved agricultural prediction. A significant challenge faced by Indian farmers is selecting the right crop for their specific soil conditions. To overcome his, [1] developed a system for crop classification that employs deep reinforcement learning (DRL) to support precision agriculture by recommending suitable crops for farmers. This system was evaluated against various machine learning techniques, including K-Nearest Neighbor (KNN), Naive Bayes, and Random Forest, to ensure both high accuracy and efficiency in suggesting crops tailored to specific locations.

2.3 Ensemble Learning in Agriculture

Precision agriculture, a contemporary farming approach, seeks to address this challenge for farmers. Recent progress in deep learning has revealed its capacity to surpass traditional statistical models in crop forecasting. While ensemble methods have been widely used in various fields, they are only now gaining traction in crop prediction due to enhancements in deep learning frameworks and processing power. Deep learning architectures, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short Term Memory Networks (LSTMs), have shown the capability to recognize both spatial and temporal relationships in agricultural datasets. By combining these models, ensemble strategies such as stacking, bagging, and boosting have the potential to improve predictive accuracy. A study [18] reviewed deep learning architectures such as CNN and LSTM for crop yield prediction and highlighted their effectiveness in modelling complex agricultural datasets. Another study [20] recommended merging of different deep learning frameworks, including long short term memory (LSTM) networks and convolutional neural networks (CNNs), for forecasting crop yields based on information obtained from remote sensing. The findings indicate that the ensemble approach outperforms standalone models in terms of both precision and resilience, showcasing the potential of deep learning methods to enhance the accuracy of crop yield forecasts. In a similar vein, [11] reviewed recent advancements in ensemble learning and explainable artificial intelligence techniques for agricultural applications and emphasized their effectiveness in addressing challenges associated with climate variability and crop yield prediction. [8] introduced a hybrid ensemble model that fused an LSTM network with a stacking ensemble methodology for soybean yield predictions. This approach not only improved prediction accuracy but also effectively captured the temporal relationships within the dataset. This literature review explores the evolution of deep learning methodologies and ensemble strategies, as well as their applications in crop and fertilizer recommendations illustrated in tabular format in Table 1. It also highlights the challenges and identifies existing research gaps in the domain.

Table 1: Review of Machine Learning and Deep Learning Approaches for Forecasting Crop Yield and Fertilizer Recommendation

Study	Methods Used	Dataset	Limitation	Contribution
[1] Bouni et al. (2022)	DRL, KNN, RF, Naive Bayes	Location specific crop data	Limited scope of DRL generalization	Crop recommendation tailored for precision agriculture using DRL
[3] Jaiswal et al. (2020)	Collaborative Filtering, ML models (general)	Agriculture sector dataset	Not tailored to soil specific recommendation	Recommender system for crop suggestions in agriculture
[6] Pande et al. (2021)	RF, SVM, MLR, ANN, KNN	Crop productivity data with GPS	Accuracy dependent on GPS and input quality	ML based mobile app for crop and fertilizer recommendation
[7] Chakraborty et al. (2021)	ML based crop recommendation	Structured data: soil type, season, location	Not scalable for complex, nonlinear data	Site specific crop recommendation system
[16] Motwani et al. (2022)	Random Forest, CNN	Soil and regional datasets	Limited scalability; RF lower accuracy (75%)	High accuracy (95.21%) with CNN in crop recommendation

Study	Methods Used	Dataset	Limitation	Contribution
[17] Logeshwaran et al. (2024)	Deep neural networks	Soil and weather dataset	High computational cost	Deep learning framework for agriculture
[21] Pruthviraj et al. (2022)	SVM, Soil Classification	Soil sample datasets	Limited to four fertility levels	Fertility classification and crop/fertilizer recommendation using SVM
[18] Wang et al. (2024)	CNN, LSTM	Agricultural crop yield datasets	Complex model structure	Deep learning crop yield prediction
[20] Li et al. (2020)	CNN + LSTM Ensemble	Remote sensing data	Model complexity and training cost	Improved accuracy and resilience in crop yield prediction
[11] Choi et al. (2025)	Ensemble learning, Deep Learning, Explainable AI techniques	Crop yield and climate datasets	High computational cost	Ensemble learning techniques for crop yield prediction under varying climate conditions
[8] Wang et al. (2020)	LSTM + Stacking Ensemble	Soybean yield data	May not generalize to other crops	Captures temporal features for improved yield prediction
[23] Goel et al. (2026)	ANN and Optimization	Remote sensing data	High complexity	Hybrid crop recommendation model
[22] Li et al. (2020)	Deep Learning Ensemble	Remote sensing	Resource intensive	Yield prediction using remote sensing + DL models
[24] Afzal et al. (2025)	Machine learning algorithms (Random forest, Decision Trees)	Soil nutrient dataset	Limited nonlinear modeling	Soil aware crop recommendation
[25] Chen et al. (2021)	Ensemble ML	Corn yield + Remote sensing	Limited to one crop type	Ensemble for corn yield prediction using satellite data
[26] Singh & Sharma (2025)	Machine Learning and Cloud computing	Real time agricultural dataset	Cloud dependency	cloud-enabled crop recommendation system
Proposed AgroDeepFusion	CNN+LSTM+DNN Ensemble	Soil laboratory dataset (2018-2023) containing macro, micro nutrients and soil properties	Region specific dataset (Tamil Nadu)	Hybrid deep ensemble architecture for crop prediction and fertilizer recommendation with improved accuracy

While prior research on crop recommendation using traditional machine learning, deep learning, and ensemble based approaches has shown promising results, several limitations remain. Traditional machine models depend on manual feature engineering and struggle to effectively model the complex interactions between soil and crops [10]. On the other hand, many deep learning architectures, particularly CNN, LSTM frameworks, are primarily designed to handle either temporal or remote sensing data, making them less suitable for soil centric datasets. Furthermore, existing ensemble methods in agriculture often combine similar learners that leads to a lack of robustness and limited statistical validation. To address these research gaps, the proposed AgroDeepFusion framework integrates CNN, LSTM, and DNN models.

2.4 Proposed Methodology: AgroDeepFusion Framework

The proposed AgroDeepFusion ensemble design framework features a layered architecture that aims to overcome the limitations of standalone models and enhance robustness, especially in handling the noisy or incomplete datasets that are often encountered in real world farming. The proposed architecture consists of several phases, including data acquisition, pre-processing, model development, and performance evaluation. By integrating CNN, LSTM, and DNN layers, the model can accurately predict optimal crops based on soil characteristics with improved accuracy.

2.5 Input Layer

2.6 Dataset Collection and Preprocessing

For this study, the dataset used was collected from government and private agricultural soil test laboratories situated across Tamil Nadu, India. It covers five years of data from 2018 to 2023, representing a diverse range of soil samples and farming conditions. The dataset includes soil characteristics, such as pH, EC, OC, and both macro and micro nutrients such as N, P, K, Zn, Fe, B, and more. Each soil sample is associated with a target crop class, enabling supervised learning across 22 different crop categories.

The dataset used in this study represented as

$$D = \{(x_i, y_i)\}_{I=1}^{N} \qquad (1)$$

Where xi represents the soil feature vector for the ith sample and yi denotes the corresponding crop label. Although this region specific dataset is based observations from soil labs in Tamil Nadu, the regional and temporal diversity enhances the model's strength by promoting generalization across different field conditions. Additionally, the proposed ensemble learning methodology is designed to be generic and data driven, making it applicable to soil based datasets from other geographic regions, provided appropriate re training and calibration are conducted using region specific data. Sample datasets are provided in the supplementary materials for reference shown in Table 2 [19].

Table 2: Features of Soil Dataset

1	**Macro Nutrients**	**Nitrogen (N)**
2	Macro Nutrients	Phosphorus (P)
3	Macro Nutrients	Potassium (K)
4	Macro Nutrients	Sulfur (S)
5	Micro Nutrients	Zinc (Zn)
6	Micro Nutrients	Iron (Fe)
7	Micro Nutrients	Copper (Cu)
8	Micro Nutrients	Manganese (Mn)
9	Micro Nutrients	Boron (B)
10	Physical and Chemical Properties	Soil pH
11	Physical and Chemical Properties	Electrical Conductivity (EC)
12	Physical and Chemical Properties	Organic Carbon (OC)
13	Physical and Chemical Properties	L.S
14	Physical and Chemical Properties	TEX
15	Physical and Chemical Properties	Soil Colour
16	Target Variable	Crop

Data Preprocessing consists of four steps data cleaning, feature normalization, categorical feature encoding and dataset partitioning. The collected dataset was examined for missing values and it was handled using interpolation and imputation techniques to preserve the integrity of the dataset. The soil dataset contains attributes which is measured in different numerical ranges. MinMax normalization was applied to transform the all input variables into uniform range between 0 and 1. The normalization process is defined as

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (2)$$

where x represents the original feature, xmin and xmax represents the minimum and maximum values of the feature within the dataset. This transformation is helpful to ensure that all the features contribute equally to the learning process. In addition to the numerical attributes, the dataset contains categorical features such as soil color, texture and crop classification. These categorical features were converted as numerical features using one hot encoding. The preprocessed soil attributes are organized into structured feature groups representing progressive soil characteristics. These feature groups are arranged in a specific order rather than as independent variables. This arrangement enables the model to capture cumulative and interdependent soil fertility relationships, instead of treating each soil property as an isolated variable. After preprocessing, the dataset was partitioned using stratified splitting strategy to ensure a representative distribution of crop classes across training and testing subsets. Specifically, 80% of the dataset samples were allocated for training and the remaining 20% for testing.

2.7 Processing Layer: AgroDeepFusion Hybrid Architecture

We developed AgroDeepFusion, a hybrid ensemble model that integrates CNN, LSTM, and DNN architectures. The proposed framework integrates CNN, LSTM, and fully connected deep neural networks (DNNs) using stacking and bagging techniques in a single pipeline that was depicted in Figure 2 and the Pseudo Code for Proposed Ensemble Model (AgroDeepFusion) is depicted in Figure 3.

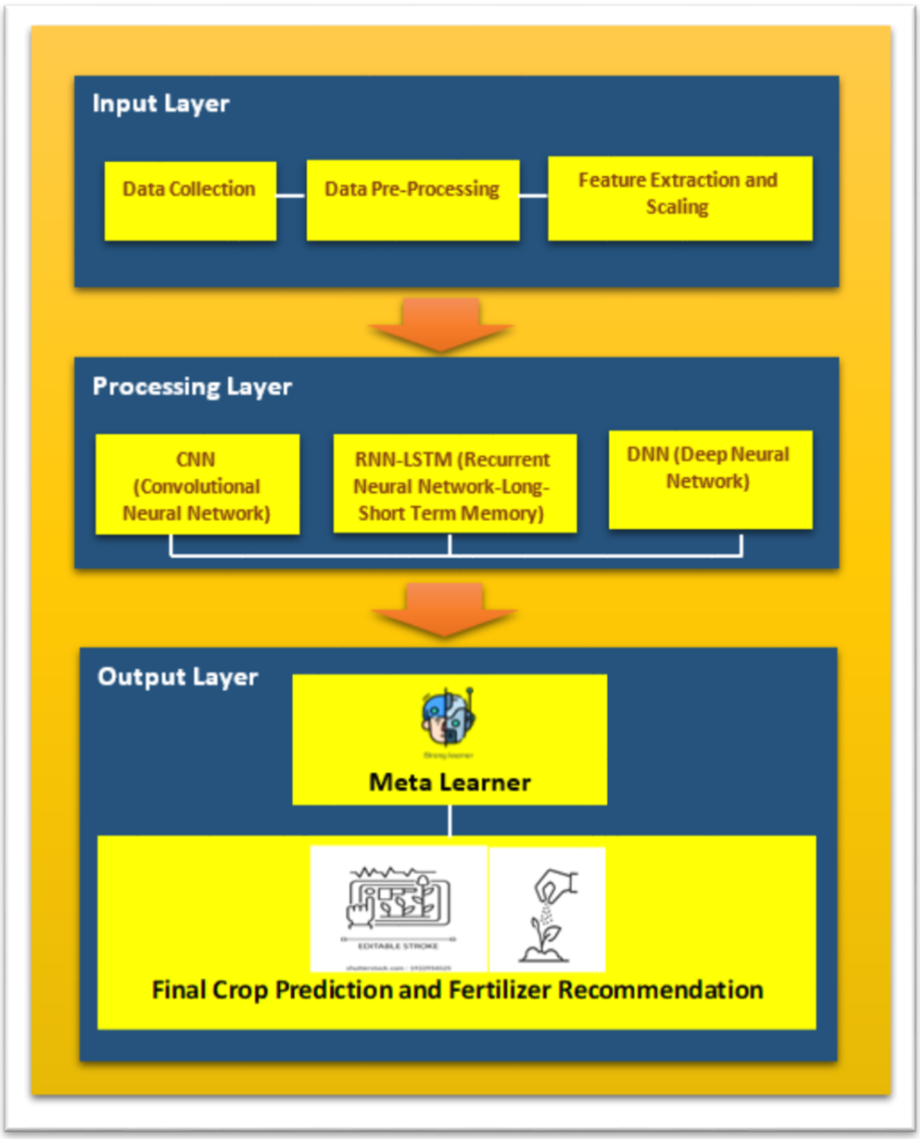

Figure 2: AgroDeepFusion: Proposed Ensemble Model Architecture

```
BEGIN
    # Step 1: Data Collection
    # Step 2: Feature Engineering
        1. Remove unnecessary columns that are not useful for prediction.
        2. Identify the target column (which is the type of crop we want to predict).
           Identify Numbers and Categories columns.
        3. Handle missing values. Identify the important features using confusion
           matrix and perform correlation analysis.
        4. View the correlation analysis by using a visual graph (heatmap) to show
           relationships.
    # Step 3: Data Transformation
        1. Convert category-type data into numbers. Scale all numerical data to make
           sure all values are in the same range.
        2. Convert the crop column (target column) into numbers. Identify the target
           column(Crop) with numerical values.
    # Step 4: Split Data into Training and Validation Sets
        1. Separate the input data (soil properties) and target data (crop type).
        2. Divide the dataset into 80% for training (to learn patterns) and 20% for
           testing (to check accuracy).
    # Step 5: Create Deep Learning Models
        1. Model 1: Convolutional Neural Network (CNN)
        2. Model 2: Recurrent Neural Network (RNN - LSTM)
        3. Model 3: Deep Neural Networks (DNN)
    # Step 6: Combine Models and create Ensemble Model (AgroDeepFusion)
        1. Take the outputs of MLP, CNN, and RNN models.
        2. Combine them together into one Ensemble Model (a super model).
        3. Add a final softmax layer that makes the meta learner for final crop
           prediction.
    # Step 7: Train the Model
      1. Prepare the model for learning using Adam optimizer (a tool that improves
         accuracy).
      2. Train the model for 50 rounds (epochs) using small batches of 64 samples at a
         time.
    # Step 8: Evaluate Model Performance
      1. Test the trained model on the 20% testing data that was set aside earlier.
      2. Measure how well it predicts the correct crop.
    # Step 9: Predict Crop Based on User's Soil Data
      1. Allow users to enter details about their soil (such as pH, temperature, and
         moisture).
      2. Use the trained model to predict the best crop for the given soil.
      3. Display the recommended crop to the user.
    # Step 10: Recommend Fertilizer
      1. Recommend Fertilizer using threshold-based logic on NPK values to suggest
         nutrient-specific fertilizers tailored to soil deficiencies.
END
```

Figure 3: Pseudo Code for Proposed Ensemble Model(AgroDeepFusion)

CNN architecture consists of Conv1D layer with 64 filters and kernel size of 3 followed by ReLU activation function, a max pooling layer, and a dropout layer with a rate of 0.2 to prevent overfitting. CNN is employed to extract abstract spatial characteristics from the input feature vector through 1D convolution. The convolution operation is defined as

$$h_i = f(W * x_i + b) \tag{3}$$

Where xi represents the input vector, W denotes the convolution kernel, b is the bias and f is the activation function. This process helps to identify significant combinations of soil and weather variables that affect crop development [25]. The LSTM network captures temporal dependencies and contextual details derived from the CNN's output. The internal operations of the LSTM are structured by gating mechanisms that regulate the flow of information. The forgot gate is defined as

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \quad (4)$$

And the Input gate is defined as

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (5)$$

Where xt represents the input at time step t, ht-1 denotes the previous hidden state, Wf and Wi represents weight matrices, bf and bi denotes the bias terms. The LSTM layer consists of 64 memory units with 0.2 dropout rate to reduce overfitting. This LSTM enables the model to capture hidden dependencies among soil properties such as texture, nutrient availability and physicochemical characteristics which influences the crop suitability. Simultaneously, a fully connected DNN processes additional features like soil type or lime status, to model complex, non-linear interactions. The DNN architecture comprises of two full connected layers with 128 and 64 neurons followed by ReLU activation function. The transformation in each dense layer is expressed as

$$z = f(W_x + b) \quad (6)$$

Where x represents the input feature vector, W denotes weight matrix, b is the bias term and f is the activation function. The outputs from all three models integrated using a stacking based ensemble strategy and fed into a meta learner, typically a logistic regression classifier that synthesizes the predictions and makes the final crop recommendation [22]. The ensemble prediction can be expressed as

$$\mathrm{P_{final}} = \mathrm{f_{meta}}\,(\mathrm{P_{cnn}}, \mathrm{P_{lstm}}, \mathrm{P_{dnn}}) \quad (7)$$

Where Pcnn , Plstm , and Pdnn represents the outputs of the CNN, LSTM, and DNN models respectively, and fmeta represents the meta learning implemented using logistic regression classifier. This stacking based ensemble strategy strengthens the model and enabling the framework to capture spatial patterns, sequential relationships, and nonlinear feature interactions within the soil dataset.

2.8 Output Layer: Crop Prediction and Fertilizer Recommendation

The final ensemble output generates two key decision support outcomes i.e. crop prediction and fertilizer recommendation. Crop prediction, which identifies the most suitable crops based on the soil characteristics and fertilizer recommendation, which utilizes the predicted crop along with the soil nutrient levels to suggest appropriate fertilizer types and application levels through a rule based agronomic module.

2.9 Experimental Results

To assess the effectiveness and robustness of the proposed hybrid ensemble model, various performance indicators as defined in Eq. (8) to Eq. (11) were employed. The evaluation outcomes for these metrics are summarized in Table 3.

Table 3: Performance Metrics for Evaluation

S. No.	Metrics for Evaluation
1	Accuracy
2	Precision
3	Recall
4	F1-Score

Accuracy: Accuracy evaluates model correctness by determining the ration of accurate crop predictions to the total predictions generated.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \quad (8)$$

Precision: Precision, also referred to as the Positive Predictive Value, measures the proportion of correctly identified instances among all samples predicted as a specific crop class (e.g., "Wheat").

$$\text{Precision} = \frac{\text{True Positives(TP)}}{\text{True Positives(TP)+False Positives(FP)}} \quad (9)$$

Recall: Recall, or Sensitivity, evaluates the model's capability to correctly identify all relevant instances belonging to a particular class.

$$\text{Recall} = \frac{\text{True Positives(TP)}}{\text{True Positives(TP)+False Negatives(FN)}} \quad (10)$$

F1-Score: The F1-score combines precision and recall through their harmonic mean, offering a balanced measure of the model's performance for ensuring both false positives and false negatives are equally accounted when dealing with skewed or imbalanced datasets.

$$\text{F1} - \text{Score} = 2 * \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

The performance of the proposed ensemble model was assessed using the metrics such as accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrix, demonstrated a 95% accuracy that outperforms traditional machine learning and standalone deep learning models. The confusion matrix revealed minimal false positives, crucial for accurate fertilizer recommendations. The model also demonstrated high computational efficiency, with a training time of 18 minutes on a GTX 1660 Ti GPU and an inference time of just 0.003 seconds per sample, making it suitable for real time applications on mobile or embedded devices. Robustness testing with noise injected data showed only a 2% drop in accuracy, indicating strong resilience to incomplete or imperfect input. Additionally, a user friendly web interface built with Streamlit allows users to input soil data and receive crop and NPK fertilizer recommendations, enhancing real world applicability. These results confirm the model's strong predictive capabilities, low complexity, and adaptability, making it a practical and scalable solution for smart agriculture and decision making support in diverse farming conditions.

2.10 Performance Comparison with Traditional Models

Traditional regression models and Support Vector Machines (SVMs) have shown moderate success in crop prediction tasks but often struggle with high-dimensional agricultural datasets that involve complex interdependencies between variables [21][16][17]. In contrast, ensemble deep learning approaches have demonstrated superior performance. Prior studies, such as those by Choi et al. [11] and Wang et al. [18], showed that hybrid CNN, LSTM models can outperform standalone networks by efficiently learning spatial patterns and temporal dynamics from soil and environmental data. However, these models often lack robustness and generalizability across diverse agro climatic zones and incomplete datasets. Hybrid stacking approaches that incorporate boosting techniques, as seen in the work by Zhang et al. [20], have improved scalability and accuracy but remain constrained by dataset specific limitations. The proposed model, AgroDeepFusion, addresses these gaps by integrating CNN, LSTM, and DNN into a unified ensemble. This architecture leverages spatial, sequential, and nonlinear feature learning to enhance prediction quality. The proposed model achieved a validation accuracy of 95.0%, significantly surpassing the standalone CNN (86.5%), LSTM (87.5%), and DNN (84.0%) models, with a 15% gain in F1-score and around 10% improvement in both precision and recall. These results, detailed in

Table 4 and depicted in Figure 4 and Figure 5, confirming its robustness and general applicability.

Table 4: Performance Comparison with traditional models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Decision Tree	78.0	76.0	75.0	75.5
SVM	82.5	80.0	79.0	79.5
Random Forest	85.0	83.5	82.5	83.5

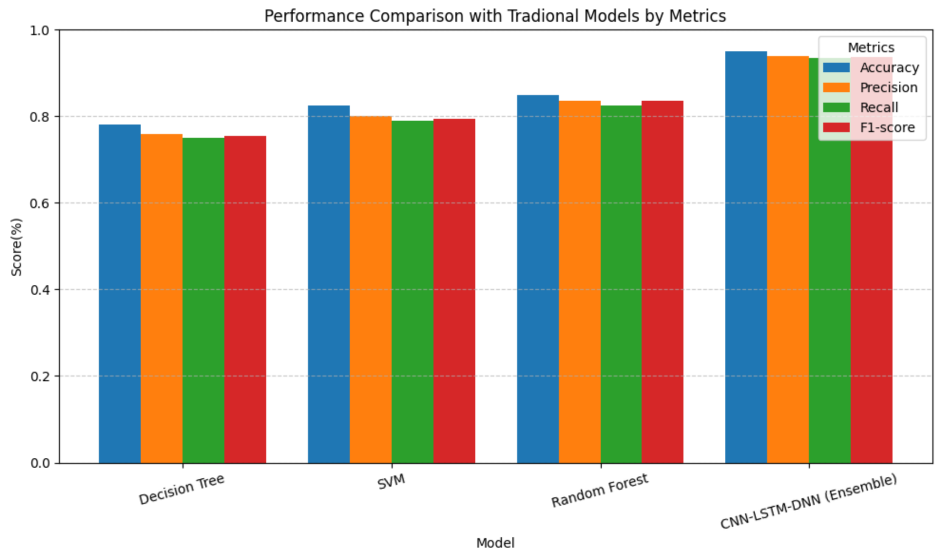

Figure 4: Performance Comparison with Traditional Models

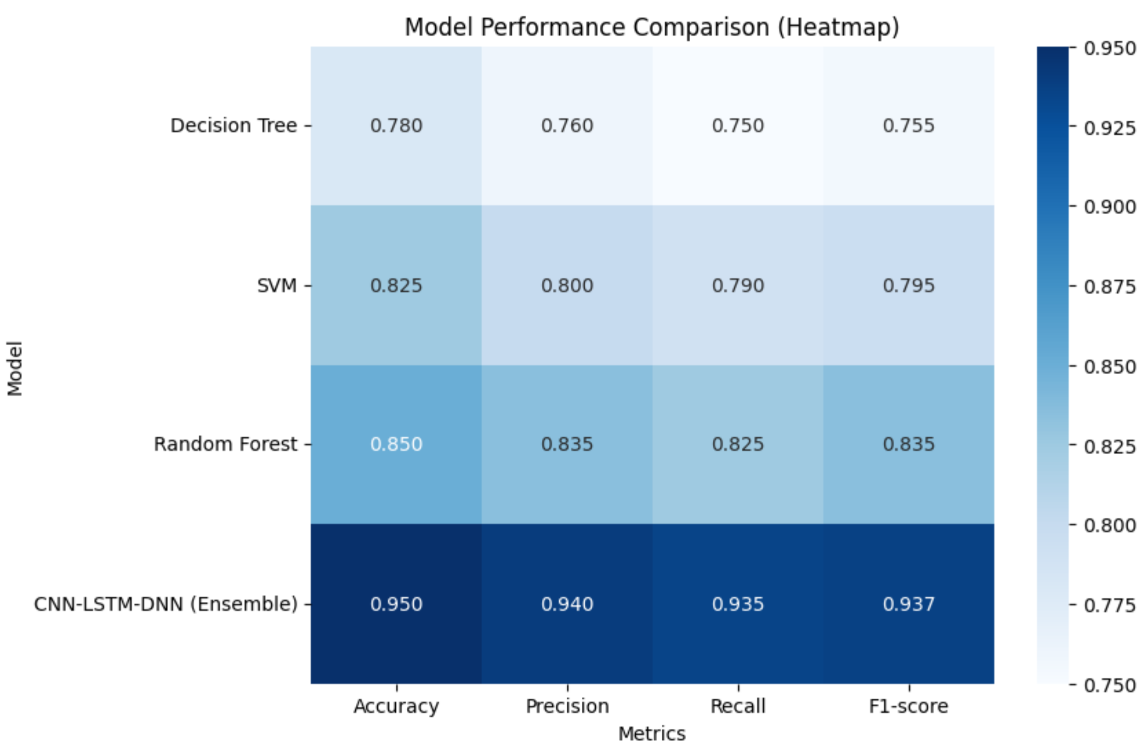

Figure 5: Performance Comparison with Traditional Models (HeatMap)

3. Discussion

The stacked ensemble model combining CNN, LSTM, and DNN demonstrated significant improvements in prediction accuracy, achieving a validation accuracy of 95%, surpassing standalone CNN (86.5%), LSTM (87.5%), and DNN (84.0%) models. The ensemble attained an F1-score of 93.7%, precision of 94%, and recall of 93.5%, effectively capturing spatial, temporal, and nonlinear relationships within the soil dataset. These results, summarized in Table 5 and illustrated in Figure 6. underscore the ensemble's ability to model complex soil–crop interactions while mitigating overfitting by integrating the strengths of multiple architectures. The integration of a rule based fertilizer recommendation module further improves its practical utility, offering farmers data driven guidance for effective crop selection and nutrient management. Compared to traditional machine learning and individual deep learning approaches, the proposed framework exhibits superior accuracy and generalizability across varied soil conditions. However, it is important to note that the experimental evaluation is conducted using a region specific dataset derived from Tamil Nadu, India. While this regional focus improves the reliability of inference within agro ecological conditions typical of South Indian Farming systems and enables contextually relevant modelling, it also introduces a potential limitation with respect to generalizability. Without proper adaptation or retraining, the patterns learned by the model may not directly transfer to other geographical regions because they might represent geoclimatic factors, soil characteristics, and agricultural management practices specific to this region. Furthermore, integration of CNN, LSTM and DNN components, the black box nature of deep learning presents interpretability challenges. Partial interpretability has been addressed through comparative model assessment, which provide indirect insight into the contribution of structured soil feature encoding. Future work will focus on incorporating explainable AI tools such as SHAP and LIME and exploring attention mechanisms and Transformer based architectures, along with multi modal data integration, to boost both performance and usability in precision agriculture [2] [11].

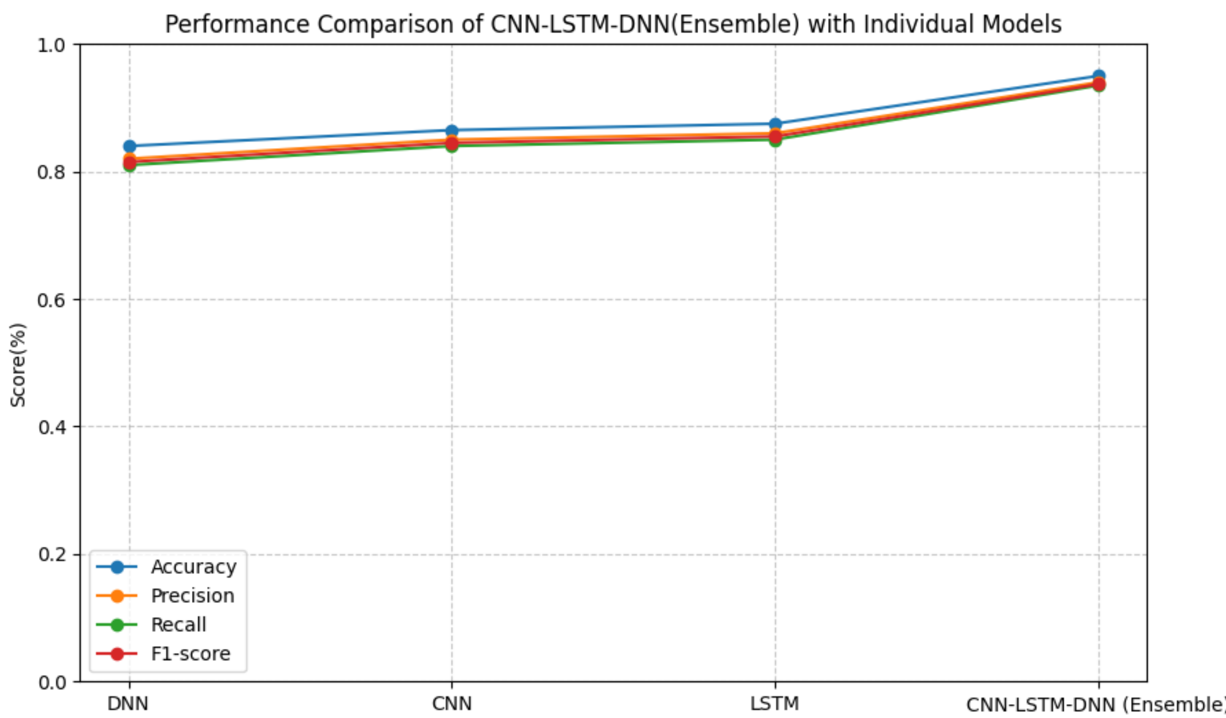

Figure 6: Performance Comparison of AgroDeepFusion Framework (CNN-LSTM-DNN) with Individual Models

Table 5: Performance Comparison of Individual models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
DNN	84.0	82.0	81.0	81.5
CNN	86.5	85.0	84.0	84.5
LSTM	87.5	86.0	85.0	85.5
AgroDeepFusion Framework (CNN-LSTM-DNN)	95.0	94.0	93.5	93.7

4. Conclusion

This research propose a hybrid ensemble deep learning framework designed for predictive analysis and fertilizer recommendation using real world test data. The proposed architecture integrates the strengths of DNN, LSTM and CNN within a stacking ensemble format. Through comprehensive preprocessing techniques including feature scaling, encoding, and managing missing values the model successfully discerned intricate connections between soil characteristics and crop viability. This study presents a hybrid deep ensemble model architecture combining DNN, LSTM, CNN for crop and fertilizer recommendation using soil lab test data from India. A key contribution of this work is the reinterpretation of the LSTM module as a sequence aware encoder for structured soil feature profiles rather than as a temporal model. By grouping soil texture, chemical, and nutrient attributes into ordered feature sequences, the model captures cumulative and inter dependent fertility relationships that influence crop productivity. The ensemble achieved a notable accuracy of 95% and surpassed existing machine learning and deep learning baselines in recall, F1-score and precision as shown in Table 5 and depicted in Figure 6. By integrating a rule based fertilizer recommendation and evaluating the model's robustness and scalability, this work provides a step toward practical precision agriculture solutions. Future studies will focus on broadening the model's capabilities by integrating transformer architectures and attention mechanisms and applying explainable artificial intelligence techniques such as SHAP and LIME to improve model interpretability, which have shown promise in various domains. Investigating advanced ensemble strategies such as adaptive boosting and weighted stacking may further improve prediction generalization and precision. Expanding the dataset that will involve real time integration with mobile applications and deploying the model in farming communities across multiple Indian states. These advancements will enable precision agriculture solutions not only more reliable and efficient but also environmentally sustainable.

4.1 Abbreviations

ML: Machine Learning, DL: Deep Learning, CNNs: Convolutional Neural Networks, LSTMs: Long Short-Term Memory networks, DNNs: Deep Neural Networks, DRL: Deep Reinforcement Learning, SVM: Support Vector Machine, KNN: K-Nearest Neighbour, MLR: Multivariate Linear Regression, ANN: Artificial Neural Network, SHAP: SHapley Additive exPlanations and LIME: Local Interpretable Model agnostic Explanation.

Acknowledgement

None.

Author contributions

Sankareswari K: Data Collection, Data Analysis, Conceptualization, Model Design, Implementation, Testing and wrote the main manuscript text. G.Sujatha: Model Design and offered critical suggestions on the manuscript text.

Conflict of Interest

The authors declare no conflict of interest related to this research work, financial or otherwise.

4.2 Ethics approval

This study did not involve any human or animal subjects. Informed consent was not required for this study as it did not involve human participants.

4.3 Availability of data and materials

Availability of data and materials: The datasets generated and/or analysed during the current study are not publicly available due to its collected from Soil test lab for corresponding author Ph.D work but are available from the corresponding author on reasonable request

Funding

Authors states that this study was conducted without any financial support.

Reference

1. Bouni M, Hssina B, Douzi K, & Douzi S (2022) Towards an Efficient Recommender Systems in Smart Agriculture, A deep reinforcement learning approach. Procedia Computer Science, 203, 825–830. doi: 10.1016/j.procs.2022.07.124
2. Shastri R, Patel K, Singh D (2025) Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence. Scientific Reports. https://doi.org/10.1038/s41598-025-07003-8
3. Jaiswal, Sapna, Tejaswi Kharade, Nikita Kotambe, and Shilpa Shinde (2020) Collaborative Recommendation System For Agriculture Sector. In ITM Web of Conferences, vol. 32. EDP Sciences, DOI: https://doi.org/10.1051/itmconf/20203203034.
4. Suresh, Dr & Kumar, Dr. (2021). Efficient Crop Yield Recommendation System Using Machine Learning For Digital Farming.

5. Kumar, Y. Jeevan Nagendra, V. Spandana, V. S. Vaishnavi, K. Neha, and V. G. R. R. Devi (2020). Supervised Machine Learning Approach for Crop Yield Prediction in Agriculture Sector. In 2020 5th International Conference on Communication and Electronics Systems (ICCES),pp.736-741.IEEE. 10.1109/ICCES48766.2020.9137868.
6. Pande S M, Ramesh P K, Anmol A, Aishwarya B R, Rohilla K and Shaurya K (2021) Crop Recommender System Using Machine Learning Approach, 5th International Conference on Computing Methodologies and Communication (ICCMC), pp. 1066-1071, doi: 10.1109/ICCMC51019.2021.9418351
7. Chakraborty P A S, Kumar A and Pooniwala O R (2021) Intelligent Crop Recommendation System using Machine Learning," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), pp. 843-848, doi: 10.1109/ICCMC51019.2021.9418375
8. Wang, P., Zheng, L., Wang, X., Xie, X., Zhao, Y., & Wang, C. (2020). Ensemble-based models for crop yield prediction: A case study on wheat. Computers and Electronics in Agriculture, 178, 105751.
9. Hussain W, Mushtaq M F, Shahroz M et al. (2025) Ensemble genetic and CNN model-based image classification by enhancing hyperparameter tuning. Scientific Reports, 15, 1003. https://doi.org/10.1038/s41598-024-76178-3.
10. Siddiqi M H, Khan A, Lee S (2023) Multi-criteria agriculture recommendation system using machine learning for crop and fertilizer prediction. Current Agriculture Research Journal, 11(1), 75-86.
11. Choi J W, Hidayat M S, Cho S B, Hwang W H, Lee H, Kim M S (2025) Recent trends in machine learning, deep learning, ensemble learning, and explainable artificial intelligence techniques for evaluating crop yields under abnormal climate conditions. Plants, 14(18), 2841. https://doi.org/10.3390/plants14182841.
12. Li, Q., Ye, Z., & Chen, J. (2020). Crop yield prediction based on hybrid ensemble learning methods. Information Processing in Agriculture, 7(4), 476-484.
13. Hasan M, Marjan M A, Uddin M P, Afjal M I, Kadry S, Nam Y (2023) Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop cultivation. Frontiers in Plant Science, 14, 1234555. https://doi.org/10.3389/fpls.2023.1234555.
14. Perwej Y, Siddiqui F, Rizvi H, Akhtar N (2025) Modelling crop yield with deep CNN-LSTM for spatiotemporal data analysis. International Journal of Scientific Research in Computer Science Engineering and Information Technology, 11(5), 54-64.
15. Rehman, S., Asif, M., Shahzad, M., et al., "Ensemble models for crop yield prediction using satellite imagery", International Journal of Agricultural and Biological Engineering, 2019.
16. Motwani A, Patil P, Nagaria V, Verma and Ghane S (2022) Soil Analysis and Crop Recommendation using Machine Learning, International Conference for Advancement in Technology (ICONAT) pp. 1-7, doi: 10.1109/ICONAT53423.2022.9725901.
17. Logeshwaran P, Kumar V, Sharma R (2024) Improving crop production using an agro-deep learning framework in precision agriculture. BMC Bioinformatics, 25, 350. https://doi.org/10.1186/s12859-024-05970-9.
18. Wang Y, Liu Z, Zhang H (2024) Progress in research on deep learning-based crop yield prediction. Agronomy, 14(10), 2264. https://doi.org/10.3390/agronomy14102264.
19. Zhang, L., Liu, J., Zhai, S., & Li, X. (2019). An ensemble learning-based method for crop yield prediction using MODIS NDVI time series. Journal of Applied Remote Sensing, 13(4), 046505.
20. Li, L., Yang, J., Huang, M., & Hu, B. (2020). Ensemble modeling for maize yield prediction: A comparison study. PLoS ONE, 15(7), e0236207.
21. Pruthviraj, Akshatha, Shastry G C, Nagaraj K A, Nikhil (2022). Crop and Fertilizer Recommendation System Based on Soil Classification. In: Shetty D., P., Shetty, S. (eds) Recent Advances in Artificial Intelligence and Data Engineering. Advances in Intelligent Systems and Computing, vol 1386. Springer, Singapore. Doi: 10.1007/978-981-16-3342-3_3
22. Li, Z., Ma, J., Liang, Y., et al. "Ensemble of deep learning models for crop yield prediction using remote sensing data", Remote Sensing, 2020.
23. Goel P, Gupta R, Sharma N (2026) Improving precision agriculture using integrated bio-inspired optimization models for crop recommendation. Scientific Reports. https://doi.org/10.1038/s41598-026-37863-7.
24. Afzal H, Khan A, Ahmed S, et al. (2025) Incorporating soil information with machine learning for crop recommendation to improve agricultural output. Scientific Reports. https://doi.org/10.1038/s41598-025-88676-z.
25. Chen, J., Wu, J., Li, J., & Chen, C. (2021). Ensemble machine learning-based approach for corn yield prediction using remote sensing data. Journal of Applied Remote Sensing, 15(1), 016508.
26. Singh G, Sharma S (2025) Enhancing precision agriculture through a cloud-based transformative crop recommendation model. Scientific Reports. https://doi.org/10.1038/s41598-025-93417-3.